

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: May 15, 2002, 09:48:20 ; Search time 19.89 Seconds  
(without alignments)  
391.314 Million cell updates/sec

Title: US-09-854-864-13

Perfect score: 498

Sequence: 1 CS0NEYFDSLHACIPCOLR.....SEYFDSLHACPRATCPQRC 81

Scoring table: BIOSYM62

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	201	40.4	184	2	B-cell maturation
2	94.5	19.0	223	2	ultra-high-sulfur
3	94.5	19.0	230	2	ultra-high-sulfur
4	92.5	18.6	169	1	ultra-high-sulfur
5	91.5	18.4	186	2	ultra-high-sulfur
6	89	17.9	1574	2	MEGF6 protein - ra
7	87	17.5	1680	2	furin (EC 3.4.21.7
8	86.5	17.4	1548	2	serine proteinase
9	86	17.3	188	2	high sulfur protei
10	83	16.7	131	1	keratin high-sulfu
11	83	16.7	175	2	keratin high-sulfu
12	82	16.5	126	2	keratin high-sulfu
13	82	16.5	526	2	cysteine-rich hair
14	81	16.3	2823	2	deubiquitinating e
15	81	16.3	2823	2	hypothetical prote
16	81	16.3	3102	2	protein T22A3.8 (1
17	80.5	16.2	151	2	hair keratin cyste
18	80.5	16.2	937	2	hair keratin cyste
19	80.5	16.2	1101	2	gene PACE4 protei
20	80.5	16.2	1798	2	hypothetical prote
21	80	16.1	3075	2	laminin beta-2 cha
22	79.5	16.0	965	2	laminin alpha-1 ch
23	79.5	16.0	1077	2	hypothetical prote
24	79.5	16.0	3133	2	probable cysteine-
25	79	15.9	861	2	hemocytin - silkw
26	78.5	15.8	201	2	Notch homolog Mote
27	78.5	15.8	294	2	hypothetical prote
28	78	15.7	572	2	hypothetical prote
29	78	15.7	1188	2	protein F14N23.5 (

30	77.5	15.6	132	1	KRCN3T	keratin high-sulfu
31	77.5	15.6	177	2	S37650	high-sulfur kerati
32	77	15.5	162	2	I47107	high-sulfur wool m
33	77	15.5	172	1	KRSHHA	keratin high-sulfu
34	77	15.5	5376	2	T42215	zonadhesin - mouse
35	76.5	15.4	932	2	I52527	PACB4A - mouse (fr
36	76.5	15.4	1513	2	T23681	hypothetical prote
37	76.5	15.4	2414	2	A54277	transcription adap
38	76	15.3	172	2	I47106	high-sulfur wool m
39	76	15.3	182	2	I47105	high-sulfur wool m
40	76	15.3	1299	2	T43251	furin (EC 3.4.21.7
41	76	15.3	2440	2	S39162	transcription coac
42	76	15.3	2441	2	S39161	CREB-binding prote
43	76	15.3	2910	2	T42214	otogelin - mouse
44	75.5	15.2	654	2	T30136	hypothetical prote
45	75.5	15.2	1713	2	A53347	adhesive ligand ep

## ALIGNMENTS

RESULT 1

S43486

B-cell maturation factor - human

N:Alternate names: BCM protein; BCMA protein; BEL protein

C:Species: Homo sapiens (man)

C>Date: 27-Jan-1995 #sequence\_revision 27-Jan-1995 #text\_change 21-Jul-2000

C:Accession: S43486; S31208; S36661

R:Laabi, Y.; Gras, M.P.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapis, A.

Nucleic Acids Res. 22, 1147-1154, 1994

A>Title: The BCM gene, preferentially expressed during B lymphoid maturation, is bid

A:Reference number: S43486; M0ID:94218235

A:Accession: S43486

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-184 <LAA>

A:Cross-references: EMBL:229574; NID:9471244; PIDN:CAA82690.1; PID:9471245

R:Laabi, Y.; Gras, M.P.; Carbonnel, F.; Brouet, J.C.; Berger, R.; Larsen, C.J.; Tsapi

EMBO J. 11, 3897-3904, 1992

A>Title: A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t(

A:Reference number: S31208; M0ID:93010984

A:Accession: S31208

A:Molecule type: mRNA

A:Residues: 1-184 <LAA>

A:Cross-references: EMBL:214954; NID:929407; PIDN:CAA78679.1; PID:929408

A:Accession: S36661

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 4-184 <LAA>

A:Cross-references: EMBL:214955

C:genetics:

A:Gene: GDB:BCMA

A:Cross-references: GDB:135977; OMIM:109545

A:Map position: 16p13.1-16p13.1

A:Introns: 44/1; 93/71

C:Superfamily: human B-cell maturation factor

Query Match 40.4%; Score 201; DB 2; Length 184;

Best Local Similarity 100.0%; Pred. No. 8.1e-11;

Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CS0NEYFDSLHACIPCOLRCSNTPTPCQRC 34

DB 8 CS0NEYFDSLHACIPCOLRCSNTPTPCQRC 41

RESULT 2

B38346

ultra-high-sulfur keratin 2 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change 24-Sep-1999

C:Accession: A38660; B38346

R.Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogell, G.  
J. Biol. Chem. 266, 4024, 1991  
A.Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin c  
A.Reference number: A38660; MUID:91154184  
A.Accession: A38660  
A.Molecule type: DNA  
A.Residues: 1-223 <MO2>  
A.Cross-references: GS:M37760; NID:g200963; PIDN:AAA40107.1; PID:g200964  
A.Note: this is a correction  
R.Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogell, G.  
J. Biol. Chem. 265, 21375-21380, 1990  
A.Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin c  
A.Reference number: A38346; MUID:91065960  
A.Accession: B38346  
A.Molecule type: DNA  
A.Residues: 1-21, 'GGCGSGGGCGSGCGGSGSCKRVC', 22-40, 'GSS', 44-45, 'G', 47-48, 'S', 50, 'GSS',  
<MO2>  
A.Cross-references: GS:M37759; NID:g200961; PIDN:AAA40106.1; PID:g200962  
A.Note: The sequence reported in this paper has been corrected. See A38660  
C:Superfamily: ultra-high-sulfur keratin

	Query Match	19.0%;	Score 94.5;	DB 2,	Length 223;
	Best Local Similarity	29.6%;	Pred. No. 0.15;		
Matches	24; Conservative			3; Indels	27; Gaps
OY	1 CSONEYFDSLHACIPCOLRCSSNRPILICORVCCCFPSLSLHACPCLCSPPTCOYCFF	60			
Dd	150 CCQGS-----CCRPC-----CGSGGCSGCCSQSCCK-----PCC-CQGSKCPCCC	189			
OY	61 HSEYFDSLHACPPATCOPYC*	81			
Dd	190 QSS-----CCRPCCGSQC	203			

```

RESULT      3
A38346
ultra-high-sulfur keratin 1 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 28-Jun-1991 #sequence_revision 28-Jun-1991 #text_change 24-Sep-1999
C:Accession: A38346
R:Wood, L.; Mills, M.; Hatzenbuehler, N.; Vogell, G.
J. Biol. Chem. 265, 21375-21380, 1990
A:Title: Serine-rich ultra high sulfur protein gene expression in murine hair and skin c
A:Reference number: A38346; MUID:91065960
A:Accession: A38346
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-230 <MOO>
A:Cross-references: GB:M37759; NID:g200961; PIDN:AAA40106.1; PID:g200962
C:Superfamily: ultra-high-sulfur keratin

```

```

Query Match          19.0%:  Score 94.5:  DB 2,  Length 230:
Best Local Similarity 29.6%:  Pred. NO.0.15:
Matches 24;  Conservative 3;  Mismatches 27;  Indels 27;  Gaps 5.

OY 1 CSONEYFDSLHACIPQCRCSNTPPLTCQRYCCYEFDSLHACPLCRSPPTQYCCF 60
    | | : | | | | | | | | | | | | | | | | | | | | | | | |
DB 149 CQGS-----CCKPC---CSSGGSSCCQSSCK-----PCC-CQSSCKPCC 188
    | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 61 HSEYFDSLHACPPATQPYC 81
    | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 189 QSS-----CCKPCCQSSC 202
    | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
SI8946
ultra high-sulfur keratin 1 - human
N:Alternate names: UHS keratin; ultra high-sulfur matrix protein
C:Species: Homo sapiens (man)
C:Date: 02-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999
C:Accession: SI8946; B36686

```

R:Drabant, B., Doenecke, D.  
submitted to the EMBL Data Library, December 1991  
A:Description: Nucleotide sequence of a Human high-sulphur keratin cDNA.  
A:Reference number: S18946  
A:Accession: S18946  
A:Molecule type: mRNA  
A:Residues: 1-169 <DR>  
A:Cross-references: EMBL:X63755; NID:g32471; PIDN:CAA45283.1; PID:g32472  
R:MacKinnon, P.J., Powell, B.C., Rogers, G.E.  
U. Cell Biol. 111, 2587-2600, 1990  
A:Title: Structure and expression of genes for a class of cysteine-rich proteins of t  
A:Reference number: A36686; MUID:91115951  
A:Accession: B36686  
A:Molecule type: DNA  
A:Residues: 1-59, 'Y', 41-169 <MAC>  
A:Cross-references: GB:X5293; NID:g34078; PIDN:CAA39005.1.; PID:g34079  
A:Genetics:  
A:Gene: GDB:KRNI  
A:Cross-references: GDB:125257; OMIM:148021  
A:Map position: 11q13-11q13  
C:Superfamily: ultra-high-sulfur keratin  
C:Keywords: hair; Tandem repeat  
F:7-15/Region: Ser-rich nonapeptide repeat  
F:59-68/Region: Gly-rich decapeptide repeat  
F:67-78/Region: Gly-rich decapeptide repeat  
F:79-88/Region: Cys-rich decapeptide repeat  
F:89-97/Region: Ser-rich nonapeptide repeat  
F:98-107/Region: Cys-rich decapeptide repeat  
F:108-117/Region: Cys-rich decapeptide repeat  
F:118-126/Region: Ser-rich nonapeptide repeat  
F:127-136/Region: Cys-rich decapeptide repeat  
F:137-145/Region: Ser-rich nonapeptide repeat  
F:146-155/Region: Cys-rich decapeptide repeat  
F:156-165/Region: Cys-rich decapeptide repeat

	Query Match	18.6%	Score 92.5	DB 1	Length 169
	Best Local Similarity	30.0%	Fred. No. 0.18		
	Matches 27	Conservative	5	Mismatches 37	Indels 21
					Gaps 5
0y	1 CSQNYFDSLHACIPQQLRCSSNPPPLTCORCCGEYFDSLHAC-PCLR-----CS 51				
Db	77 CSQ-----CSGCKPFC--CCSSGCGSCQCCSCCKPYCSQCCKPCCSSSGRSSCCQ 127				
0y	52 PPTCYCCCFHSEYFDSLHACPPATCOPYC 81				
Db	128 SSCKPCCSSSGCGSS---CCQSSCKPCC 154				

RESULT 5  
A45910 ultra-high-sulfur keratin - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 03-Jun-1993 #sequence\_revision 03-Jun-1993 #text\_change 24-Sep-1999  
C:Accession: A45910  
R:McNab, A.R.; Wood, L.; Thieriault, N.; Gierman, T.; Vogeli, G.  
J. Invest. Dermatol. 92, 263-266, 1989  
A:Title: An ultra-high sulfur keratin gene is expressed specifically during hair growth  
A:Reference number: A45910; MUID:89140394  
A:Accession: A45910  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-186 <MCN>  
A:Cross-references: GB:M27685; NID:g341749; PIDN:AAA81560.1; PID:g1066818  
A:Superfamily: ultra-high-sulfur keratin

Query Match	18.4%	Score 91.5	DB 2	Length 186
Best Local Similarity	34.7%	Pred. No. 0.24		
Matches 25: Conservative	4	Mismatches 24	Indels 19	Gaps 6

Db 85 COPCCOPSCCQSCCOPRCCSCCOPRCCISSCQPC--CRPSCQSSC-----132  
OY 70 HACPPATCQPYC 81  
Db 133 ---CRPC-CQPC 141

## RESULT 6

MEGF6 protein - rat  
T13954  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 21-Jul-2000  
C:Accession: T13954  
R:Nakayama, M.; Nakajima, D.; Nagase, T.; Nomura, N.; Seki, N.; Ohara, O.  
Genomics 51, 27-34, 1998  
A:Title: Identification of high-molecular-weight proteins with multiple EGF-like motifs  
A:Reference number: 244126; MUID:98360089  
A:Accession: T13954  
A:Status: preliminary;  
A:Molecule type: mRNA  
A:Residues: 1-1574 <NAK>  
A:Cross-references: EMBL:AB011537; NID:93449293; PIDN:BA432462.1; PID:93449294  
A:Experimental source: strain Sprague-Dawley; brain  
A:Genetics:  
A:Gene: MEGF6

Query Match 17.9%; Score 89; DB 2; Length 1574;  
Best Local Similarity 29.4%; Pred.No. 1.8;  
Matches 30; Conservative 9; Mismatches 25; Indels 38; Gaps 8;

OY 1 CSQNEFYDSLHACIPCOL-----RCSNTPPLT---CORYCEY---FDSLHACPC 47  
Db 966 CSAGAPCADAVTGSCT-CPAGRMGPRCAQSCPLTFGLNCSQICTCFNAGSCDSVTGQC-- 1022

OY 48 LRCSF-----PTCYCCFHSEFYDSLHACPPA-----TCQPYC 81  
Db 1023 -HCAPGWMGPTC-----LQACPPGLYGRNQHSC 1050

## RESULT 7

MA3434  
fruitin (EC 3.4.21.75) 2 - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 04-Mar-1993 #sequence\_revision 18-Nov-1994 #text\_change 31-Mar-2000  
C:Accession: A43434  
R:Roedbroek, A.J.; Creemers, J.W.; Pauli, I.G.; Kurzik-Dumke, U.; Rentrop, M.; Gateff, E.  
J. Biol. Chem. 267, 17408-17415, 1992  
A:Title: Cloning and functional expression of Dfruitin2, a subtilisin-like proprotein proc  
A:Reference number: A43434; MUID:92381036  
A:Accession: A43434  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1680 <ROB>  
A:Cross-references: GBI:M94375; NID:9157461; PID:9157462  
A:Note: sequence extracted from NCBI backbone (NCBIN:111933, NCBIF:111934)  
C:Genetics:  
A:Gene: FLYBase:Fur2  
A:Cross-references: FLYBase:FBgn0004598  
C:Superfamily: subtilisin homology  
C:Keywords: hydrolase; serine proteinase; transmembrane protein  
F:409-653/Domain: subtilisin homology <SBT>  
F:418-657/Domain: subtilisin homology <SBT>  
F:418-657/Active site: Asp, His, Ser #status predicted

Query Match 17.5%; Score 87; DB 2; Length 1680;  
Best Local Similarity 25.0%; Pred.No. 2.9;  
Matches 31; Conservative 12; Mismatches 31; Indels 50; Gaps 9;

OY 1 CSQNEFYDSLHACIPCOLRCSNTPPLTQCO---RYC--CEYDSLH-----ACP--- 46  
Db 1051 CPQGYEHSNRKRCVCEPCMCAS-----CQHPETCTSDH-HLVMEHKYTSACPDLT 1103

OY 47 -----CLRCSPPTCOYC--CFHSEYF--DSLHACPPA-----TC 77  
Db 1104 YETEDNKCARCHSTCATCNGPTDQDCTCRSSRYAMQNKLLISCPDGFYADKKRLCQPC 1163  
OY 78 QPYC 81  
Db 1164 QEGC 1167

## RESULT 8

S34583  
serine proteinase (EC 3.4.21.-) PC6B - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 05-Nov-1999  
C:Accession: S34583  
R:Nakagawa, T.; Murakami, K.; Nakayama, K.  
FEBS Lett. 327, 165-171, 1993  
A:Title: Identification of an isoform with an extremely large Cys-rich region of PC6,  
A:Reference number: S34583; MUID:93327934  
A:Accession: S34583  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-1548 <NAK>  
A:Cross-references: GB:D17583; NID:9407344; PIDN:BA04507.1; PID:di005033; PID:944037  
C:Keywords: hydrolase; serine proteinase

Query Match 17.4%; Score 86.5; DB 2; Length 1548;  
Best Local Similarity 22.1%; Pred.No. 3;  
Matches 29; Conservative 6; Mismatches 43; Indels 53; Gaps 3;

OY 1 CSQNEFYDSLHACIPCOLRCSNTPPLTQRCCEYFDSL----- 42  
Db 1152 CAVERWDESHRCQCHKKCSRGSESDQCTCPRETFLLMTTCYKCEPGEYHNDKDS 1211

OY 43 -----HACPRLCSP-----PRCYCCFHSEFYDSLHACP----- 73  
Db 1212 QQCVLHSSCRTEGPHSMQCLSCRPGWFLQKRECLQCHDGYGEGSTSGRCEKDKSK 1271

OY 74 -----PANCQ 78  
Db 1272 SCRGPRPTDQ 1282

## RESULT 9

JC6547  
high sulfur protein BZE - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 05-Dec-1998 #sequence\_revision 05-Dec-1998 #text\_change 21-Jul-2000  
C:Accession: JC6547  
R:Mitsui, S.; Ohuchi, A.; Adachi-Yamada, T.; Hotta, M.; Tsuboi, R.; Ogawa, H.  
Gene 208, 123-129, 1998  
A:Title: Structure and hair follicle-specific expression of genes encoding the rat hi  
A:Reference number: JC6547; MUID:98201605  
A:Accession: JC6547  
A:Molecule type: DNA  
A:Residues: 1-188 <MIT>  
A:Cross-references: DDBJ:AB003753; NID:93046870; PIDN:BA425573.1; PID:93046871  
C:Comment: This protein is a cysteine-rich, keratin associated protein.  
C:Genetics:  
A:Gene: BZE  
C:Superfamily: keratin high-sulfur matrix protein IIIA  
C:Keywords: hair

Query Match 17.3%; Score 86; DB 2; Length 188;  
Best Local Similarity 22.5%; Pred.No. 0.73;  
Matches 29; Conservative 10; Mismatches 28; Indels 62; Gaps 7;

OY 1 CSQNEFYDSLHACIP-----COLRCSNTPPLTQCO-----R 32  
Db 69 CSQSS-----CCQSCCQTSQCPPTCCONS---SCQTSCTGTGGQEGSGATSCRRV 118



Search completed: May 15, 2002, 09:49:35  
Job time: 75 sec

A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2823 <W12>  
A:Cross-references: EMBL:AL008585; PIDN:CA15432.1; GSPDB:GN00019; CESP:T22A3.8  
A:Experimental source: clone W10E24  
R:McMurray, A.  
submitted to the EMBL  
A:Reference number: Z19980  
A:Accession: T25096  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-2823 <W12>  
A:Cross-references: EMBL:Z81125; PIDN:CAB03385.1; GSPDB:GN00019; CESP:T22A3.8  
A:Experimental source: clone T22A3  
C:Genetics:  
A:Gene: CESP:T22A3.8  
A:Map position: 1  
A:introns: 45/1; 282/2; 312/3; 416/2; 1255/3; 1329/3; 1418/3; 1776/2; 1988/2; 2760/2  
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like

Query Match  
Best Local Similarity 27.7%; Score 81; DB 2; Length 2823;  
Matches 26; Conservative 10; Mismatches 32; Indels 26; Gaps 7;

QY 1 CSQNEFYDLSLHACIPQRLCSSNTPPL---TCQRYCCEYFDSLHACPLRC-----SPP 53  
DB 873 CSQGFEDPLTGKICIC--TCNGNIDPMGIGNC-----DS--ETGKCLKICGHTTGD 920  
QY 54 TCQYCCFHSEYFDSLHACPRATC-----QPYC 81  
DB 921 SCSECKEH-HWGNAQLHTCKPCGCHTQGA VNPQC 953

## RESULT 15

F87908  
Protein T22A3.8 [imported] - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 10-May-2001 #sequence\_revision 10-May-2001 #text\_change 24-Aug-2001  
C:Accession: F87908; E87908  
R:Anonymous, The C. elegans Sequencing Consortium.  
Science 282, 2012-2018, 1998  
A:title: Genome sequence of the nematode C. elegans: a platform for investigating biology  
A:Reference number: A78000; MUID:99069613; PMID:9851916  
A>Note: see websites genome.wustl.edu/gsc/\_elegans/ and www.sanger.ac.uk/Projects/C\_ele  
A:Accession: F87908  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2823 <STO>  
A:Cross-references: GB:chr\_1; PIDN:CA15432.1; PID:g3924779; GSPDB:GN00019; CESP:T22A3.8  
A:Accession: E87908  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2823 <STO>  
A:Cross-references: GB:chr\_1; PIDN:CAB03385.1; PID:g3924881; GSPDB:GN00019; CESP:T22A3.8  
C:Genetics:  
A:Gene: T22A3.8  
A:Map position: 1  
C:Superfamily: laminin alpha-1 chain; laminin G repeat homology; laminin-type EGF-like

Query Match  
Best Local Similarity 27.7%; Score 81; DB 2; Length 2823;  
Matches 26; Conservative 10; Mismatches 32; Indels 26; Gaps 7;

QY 1 CSQNEFYDLSLHACIPQRLCSSNTPPL---TCQRYCCEYFDSLHACPLRC-----SPP 53  
DB 873 CSQGFEDPLTGKICIC--TCNGNIDPMGIGNC-----DS--ETGKCLKICGHTTGD 920  
QY 54 TCQYCCFHSEYFDSLHACPRATC-----QPYC 81  
DB 921 SCSECKEH-HWGNAQLHTCKPCGCHTQGA VNPQC 953



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 09:49:40 ; Search time 10.13 Seconds  
(without alignments)  
309,604 Million cell updates/sec

Title: US-09-854-864-13

Perfect score: 496  
Sequence: 1 CSQNEFDSLHACIPQQLR.....SEYFDSLHACIPATCPYPC 81

Scoring table: BLOSUM62 ;  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	201	40.4	184	TR17_HUMAN	002223 homo sapien
2	92.5	18.6	163	KRUB_HUMAN	P26371 homo sapien
3	91	18.3	194	KRUB_HUMAN	075690 homo sapien
4	87	17.5	1680	FUR2_DROME	P30432 drosophila
5	86.5	17.4	1877	PKC5_MOUSE	004592 mus musculu
6	83	16.7	131	KRA3_SHEEP	P02441 ovis aries
7	82	16.5	526	UBPW_MOUSE	061068 mus musculu
8	81	16.3	1587	LMG3_HUMAN	094606 homo sapien
9	80.5	16.2	937	PAC4_RAT	063415 rattus norv
10	80.5	16.2	1798	LMB2_HUMAN	P55268 homo sapien
11	80	16.1	3075	LMG1_HUMAN	P25391 homo sapien
12	79.5	16.0	965	YNC3_YEAST	P53912 saccharomyc
13	79.5	16.0	3133	HMCT_BOMBO	P98092 bombyx mori
14	79	15.9	1696	PKC5_BRACL	09n115 branchiosto
15	77.5	15.6	132	KRA3_CAPII	P02442 capra hircu
16	77.5	15.6	139	KRA3_CAPII	043609 capra hircu
17	77	15.5	171	KRA2_SHEEP	P02438 ovis aries
18	77	15.5	5376	ZAN_MOUSE	088799 mus musculu
19	76.5	15.4	2414	P300_HUMAN	094672 homo sapien
20	76	15.3	2441	CBP_MOUSE	P45481 mus musculu
21	75.5	15.2	1713	LMG3_HUMAN	Q16787 mus musculu
22	75	15.1	1581	LMG3_MOUSE	Q910b6 mus musculu
23	75	15.1	1609	LMG1_HUMAN	P11047 homo sapien
24	74.5	15.0	610	CABP_BOVIN	Q28065 bos taurus
25	74	14.9	1786	LMB1_HUMAN	P07942 homo sapien
26	74	14.9	2442	CBP_HUMAN	Q92793 homo sapien
27	73.5	14.8	313	SPY1_MOUSE	Q949v9 mus musculu
28	73.5	14.8	1799	LMB2_MOUSE	Q61292 mus musculu
29	72.5	14.6	969	PAC4_HUMAN	P29122 homo sapien
30	72.5	14.6	1592	SORL1_CHICK	Q98330 g sorillin-
31	72	14.5	575	TRBM_HUMAN	P07204 homo sapien
32	72	14.5	867	SPPO_BOVIN	P98167 bos taurus
33	72	14.5	1367	IGIR_HUMAN	P08069 homo sapien

34	72	14.5	1700	BAR3_CHITE	003376 Chironomus
35	72	14.5	2569	LMG3_MOUSE	Q61789 mus musculu
36	72	14.5	4655	LRP2_HUMAN	P98164 homo sapien
37	71.5	14.4	1435	EBP1_PLARC	P19214 plasmodium
38	71	14.3	62	IBB_MEDSC	P80321 medicago sc
39	71	14.3	63	ICEL_ASCSU	P07851 ascaris su
40	71	14.3	810	NEIL_HUMAN	Q92832 homo sapien
41	71	14.3	3712	LMG_MOUSE	Q00174 drosophila
42	70.5	14.2	96	IBBA_PEA	Q41065 pisum sativ
43	70.5	14.2	1104	NFX1_HUMAN	Q12986 homo sapien
44	70	14.1	181	KR2D_SHEEP	P08131 ovis aries
45	70	14.1	3110	LMG2_HUMAN	P24043 homo sapien

## ALIGNMENTS

RESULT	ID	TR17_HUMAN	STANDARD;	PRT;	184 AA.
AC	002223;				
DT	01-JUL-1993	(Rel. 26, Created)			
DT	01-JUL-1993	(Rel. 26, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	Tumor necrosis factor receptor superfamily member 17 (B-cell maturation protein).				
GN	TNFRSF17 OR BCM4 OR BCM.				
OS	Homo sapiens (human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.				
OX	NCBI_Taxid=9606;				
RN	[1]				
RP	SEQUENCE FROM N.A., AND CHROMOSOMAL TRANSLOCATION.				
RC	TISSUE-Peripheral blood leukocytes, and lymph node;				
RX	MEDLINE=93010984; PubMed=1396583;				
RA	Laabl Y., Gras M.P., Cardonnel F., Brouet J.C.; Berger R.,				
RA	Larsen C.J., Tsapis A.;				
RT	"A new gene, BCM, on chromosome 16 is fused to the interleukin 2 gene by a t(4;16)(q26;p13) translocation in a malignant T cell lymphoma.",				
RL	EMBO J. 11:3897-3904(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=94218235; PubMed=8165126;				
RA	Laabl Y., Gras M., Brouet J., Berger R., Larsen C., Tsapis A.;				
RA	Fuhrmann J., Mason T., Crosby M.L., Barnstead M., Cronin L.,				
RA	Destaltes Mays A., Gao Y., Xu R.X., Kang H.-L., Mitchell S.,				
RT	Bichter E.E., Harris P.C., Venter J.C., Adams M.D.;				
RL	"Genome duplications and other features in 12 Mb of DNA sequence from human chromosome 16p and 16q.";				
RN	Genomics 60:295-308(1999).				
RN	[4]				
RP	SEQUENCE FROM N.A., AND VARIANT THR-153.				
RA	Kawasaki A., Tsuchiya N., Fukazawa T., Hashimoto H., Tokunaga K.;				
RT	"New Polymorphisms of human BCM4.";				
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.				
RN	[5]				
RP	FUNCTION.				
RX	MEDLINE=20363816; PubMed=10903733;				
RA	Hatzoglou A., Roussel J., Bourgade M.F., Rogier E., Madry C.,				
RA	Inoue J., Devergne O., Tsapis A.;				
RT	"TNF receptor family member BCM4 (B cell maturation) associates with				
RT	TNF receptor-associated factor (TRAF) 1, TRAF2, and TRAF3 and				
RT	activates NF-kappa B, elk-1, c-Jun N-terminal kinase, and p38				
RT	mitogen-activated protein kinase.";				
RL	J. Immunol. 165:1322-1330(2000).				
RN	[6]				

RP FUNCTION.  
RA MEDLINE-20259066; PubMed-10801128;  
RA Gross J.A., Johnston J., Mudri S., Enselman R., Dillon S.R.,  
RA Madden K., Xu W., Parrish-Novak J., Foster D., Lofton-Day C.,  
RA Moore M., Litlau A., Grossman A., Haugen H., Foley K., Blumberg H.,  
RA Harrison K., Kindsvogel W., Clegg C.H.,  
RA "TACI and BCMA are receptors for a TNF homologue implicated in B-cell  
RT autoimmunity disease";  
RL Nature 404:995-999(2000).  
CC -1- FUNCTION: Receptor for TNFSF13B/BLYS/BAFF.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN MATURE B-CELLS.  
CC -1- DISEASE: A FORM OF T-CELL ACUTE LYMPHOBLASTIC LEUKEMIA (T-ALL) IS  
CC CHARACTERIZED BY A CHROMOSOMAL TRANSLOCATION T(4;16)(Q26;P13)  
CC WHICH INVOLVES BCMA AND IL2.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z14954; CAA78679.1; -  
DR EMBL: Z29575; CAA82691.1; -  
DR EMBL: Z29574; CAA82690.1; -  
DR EMBL: U95742; AAB67251.1; -  
DR EMBL: AB052772; BAB60895.1; -  
DR PIR: S31208; S31208.  
DR PIR: S31209; S31209.  
DR MIM: 109345; -  
KW Receptor; Proto-oncogene; Transmembrane; Chromosomal translocation;  
KW Polymorphism.  
FT TRANSMEM 55 77 POTENTIAL.  
FT SITE 3 4 BREAKPOINT FOR TRANSLOCATION TO FORM  
FT VARIANT 153 153 INTERLEUKIN 2/BCM ONCOGENE.  
FT A->T.  
FT SEQUENCE 184 AA; 20138 MW; 277AF1E2767D932 CRC64;  
SQ  
Query Match 40.4%; Score 201; DB 1; Length 184;  
Best Local Similarity 100.0%; Pred. No. 4.9e-13;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CSQNEYPDSLHACIPCOLRCSSNTPPLTCORYC 34  
DB 8 CSQNEYPDSLHACIPCOLRCSSNTPPLTCORYC 41  
RESULT 2  
ID KRUH\_HUMAN STANDARD; PRT; 169 AA.  
AC P26371;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Keratin, ultra high-sulfur matrix protein A (UHS keratin A) (UHS  
DE KerA).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCB1\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-FOLLICLE;  
RX MEDLINE-91115951; PubMed-1703541;  
RA "Structure and expression of genes for a class of cysteine-rich  
RT proteins of the cuticle layers of differentiating wool and hair  
RT follicles";  
RL J. Cell Biol. 111:2587-2600(1990).

RM [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99148005; PubMed-10023043;  
RA Perez C., Auriol J., Gerst C., Bernard B.A., Egly J.-M.,  
RT "Genomic organization and promoter characterization of two human UHS  
RT keratin genes";  
RL Gene 227:137-148(1999).  
CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES  
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID  
CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-  
CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF  
CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR  
CC KERATINS (40-56 kDa).  
CC -1- TISSUE SPECIFICITY: CUTICLE LAYERS OF DIFFERENTIATING  
CC HAIR FOLLICLES.  
CC -1- DEVELOPMENTAL STAGE: AT A LATE STAGE OF FIBER DIFFERENTIATION.  
CC -1- DOMAIN: MAINLY COMPOSED OF CYS-RICH (CR), GLY-RICH (GR) AND SER-  
CC RICH (SR) REPEATS.  
CC -1- SIMILARITY: BELONGS TO THE UHS KERATIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X55293; CAA39005.1; -  
DR EMBL: AJ006693; CAA07189.1; -  
DR HSSP: P04355; 2MRT.  
DR MIM: 148021; -  
KW Keratin; Repeat; Multigene family.  
SQ SEQUENCE 169 AA; 16276 MW; 219B14FEEBA9D4A9A CRC64;  
SQ  
Query Match 18.6%; Score 92.5; DB 1; Length 169;  
Best Local Similarity 30.0%; Pred. No. 0.012;  
Matches 27; Conservative 5; Mismatches 37; Indels 21; Gaps 5;  
OY 1 CSQNEYPDSLHACIPCOLRCSSNTPPLTCORYCERFDSLHAC-PCLR-----CS 51  
DB 77 CSQ-----GSCCKKPC--CCSSGCGSSCCGCKPCGSCGCKPCSSSGRSSCCQ 127  
OY 52 PPTCYCCFHFSEYFDSLHACPPATCPQYC 81  
DB 128 SSCCKKPCSSSGCGSS---CCGSSCKKPC 154  
RESULT 3  
ID KRUH\_HUMAN STANDARD; PRT; 194 AA.  
AC O75690;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Keratin, ultra high-sulfur matrix protein B (UHS keratin B) (UHS  
DE KerB).  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCB1\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99148005; PubMed-10023043;  
RA Perez C., Auriol J., Gerst C., Bernard B.A., Egly J.-M.,  
RT "Genomic organization and promoter characterization of two human UHS  
RT keratin genes";  
RL Gene 227:137-148(1999).  
CC -1- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES  
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID  
CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-  
CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF



ID	PKCS_MOUSE	STANDARD	PRT	1877	AA
AC	004592	062040			
DT	01-FEB-1995	(Rel. 31, Created)			
DT	16-OCT-2001	(Rel. 40, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE		Protein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)			
DE		(Protein convertase PCS5) (Subtilisin/kexin-like protease PCS5)			
DE		(Convertase PCS) (PC6) (subtilisin-like protein convertase 6)			
DE		(PC6).			
GN	PCSK5				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
RP		[1]			
RP	SEQUENCE OF 330-1877 FROM N.A. (ISOFORM PCS5B).				
RC	STRAIN-ICR; TISSUE-Intestine;				
RX	MEDLINE=933379934; PubMed=8335106;				
RA	Nakayama T., Murakami K., Nakayama K.;				
RA	Nakayama K., Hosaka M., Torii S., Watanabe T., Murakami K.,				
RT	"Identification and functional expression of a new member of the				
RT	mammalian kex2-like processing endoprotease family: its striking				
RT	structural similarity to PACE4."				
RL	J. Biochem. 113:132-135(1993).				
RN		[2]			
RP	SEQUENCE FROM N.A. (ISOFORM PCS5A).				
RC	TISSUE-Brain, and Intestine;				
RX	MEDLINE=93224489; PubMed=8468318;				
RA	Nakayama T., Hosaka M., Torii S., Watanabe T., Murakami K.,				
RT	"Identification and functional expression of a new member of the				
RT	mammalian kex2-like processing endoprotease family: its striking				
RT	structural similarity to PACE4."				
RL	J. Biochem. 113:132-135(1993).				
RN		[3]			
RP	SEQUENCE FROM N.A. (ISOFORM PCS5A).				
RC	TISSUE-Adrenal cortex;				
RX	MEDLINE=93342056; PubMed=8341687;				
RA	Lusson J., Vleau D., Hamelin J., Day R., Chretien M., Seidah N.G.;				
RT	"CDNA structure of the mouse and rat subtilisin/kexin-like PCS5: a				
RT	candidate proprotein convertase expressed in endocrine and				
RT	neuroendocrine cells."				
RL	Proc. Natl. Acad. Sci. U.S.A. 90:6691-6695(1993).				
RN		[4]			
RP	PARTIAL SEQUENCE, AND SUBCELLULAR LOCATION.				
RX	MEDLINE=97103178; PubMed=8947550;				
RA	De Bie I., Marcinkiewicz M., Malide D., Lazure C., Nakayama K.,				
RA	Bendayan M., Seidah N.G.;				
RT	"The isoforms of proprotein convertase PCS5 are sorted to different				
RT	subcellular compartments."				
RL	J. Cell Biol. 135:1261-1275(1996).				
RN		[5]			
RP	DEVELOPMENTAL EXPRESSION.				
RX	MEDLINE=96293359; PubMed=8698813;				
RA	Constam D.B., Calton M., Robertson E.J.;				
RT	"SPC4, SPC6, and the novel protease SPC7 are coexpressed with bone				
RT	morphogenetic proteins at distinct sites during embryogenesis."				
RL	J. Cell Biol. 134:181-191(1996).				
RN		[6]			
RP	DEVELOPMENTAL EXPRESSION.				
RX	MEDLINE=97436919; PubMed=9291583;				
RA	Rancourt S.L., Rancourt D.E.;				
RT	"Murine subtilisin-like proteinase SPC6 is expressed during embryonic				
RT	implantation, somitogenesis, and skeletal formation."				
RL	Dev. Genet. 21:75-81(1997).				
CC	-I- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTIDASE ACTIVITY				
CC	WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE				
CC	OF CLEAVAGE AT THE RX(K/R)R CONSENSUS MOTIF. MAY BE RESPONSIBLE				
CC	FOR THE MATURATION OF GASTROINTESTINAL PEPTIDES. MAY BE INVOLVED				
CC	IN THE CELLULAR PROLIFERATION OF ADRENAL CORTEX VIA THE ACTIVATION				
CC	OF GROWTH FACTORS.				

CC -1- CATACTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR  
CC PROPOTEINS BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-2AA BONDS, WHERE XAA  
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.  
CC -1- SUBCELLULAR LOCATION: PCS5 IS SECRETED THROUGH THE REGULATED  
CC SECRETORY PATHWAY. PCS5 IS A TYPE I MEMBRANE PROTEIN LOCALIZED TO  
CC A PARANUCLEAR POST-GOLGI NETWORK COMPARTMENT IN COMMUNICATION WITH  
CC EARLY ENDOPLASM.  
CC -1- ALTERNATIVE PRODUCTS: AT LEAST 2 ISOFORMS: PCS5/LONG (SHOWN HERE)  
CC AND PCS5/SHORT: ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: PCS5 IS EXPRESSED IN MOST TISSUES BUT IS MOST  
CC ABUNDANT IN THE INTESTINE AND ADRENALS. PCS5 IS EXPRESSED IN THE  
CC INTESTINE, ADRENALS AND LUNG BUT NOT IN THE BRAIN.  
CC -1- DEVELOPMENTAL STAGE: WEAKLY EXPRESSED THROUGHOUT THE EMBRYO,  
CC EXCEPT IN THE DEVELOPING NERVOUS SYSTEM, THE RIBS AND THE LIVER,  
CC BUT MARKEDLY UPREGULATED AT DISCRETE SITES DURING DEVELOPMENT. AT  
CC E6.5, PROMINENT EXPRESSION OBSERVED IN DIFFERENTIATED DECIDUA. AT  
CC E7.5, INTENSE EXPRESSION IN EXTRAMEMBRANIC ENDODERM, ANIONIC AND  
CC NASCENT MESODERM. AT 8.5, ABUNDANT EXPRESSION IN SOMITES AND YOLK  
CC SAC FOLLOWED BY A CONFIRMATION TO DERMATOME COMPARTMENT. BETWEEN  
CC E8.5 AND E11.5, ABUNDANT EXPRESSION IN AER (THICKENED ECTODERMAL  
CC CELLS OF LIMB BUDS). AT E12.5, EXPRESSION IN THE LIMBS IS CONFINED  
CC TO THE CONDENSING MESENCHYM SURROUNDING THE CARTILAGE. AT THIS  
CC STAGE, STRONG EXPRESSION ALSO DETECTED IN VERTEBRAL AND FACIAL  
CC CARTILAGE PRIMORDIA AND IN THE MUSCLE OF THE TONGUE. AT E15.5,  
CC ABUNDANT EXPRESSION IN EPITHELIAL CELLS OF THE INTESTINAL VILLI.  
CC ISOFORM A IS MOST ABUNDANT AT ALL STAGES BUT SIGNIFICANT LEVELS OF  
CC ISOFORM B OCCUR AT E12.5.  
CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE  
CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC  
CC RETICULUM.  
CC -1- DOMAIN: AC 1 AND AC 2 (CLUSTERS OF ACIDIC AMINO ACIDS) CONTAIN  
CC SORTING INFORMATION. AC 1 DIRECTS TGN LOCALIZATION AND INTERACTS  
CC WITH THE TGN SORTING PROTEIN PACS-1.  
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE  
CC SUBTILASIN FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL, D17583; BAA04507.1; -;  
CC EMBL, D12619; BAA02143.1; -;  
CC EMBL, L14932; AAT74636.1; -;  
CC PIR: JX0248; JX0248.  
CC PIR: A48225; A48225.  
CC HSP: Q99405; IMP1.  
CC MEROPS: S08.076; -;  
CC MGD: MG1:97515; PCS5.  
CC InterPro: IPR000561; EGF-like.  
CC InterPro: IPR002174; Furin-like.  
CC InterPro: IPR002884; P\_domain.  
CC InterPro: IPR000209; peptidase\_S8.  
CC Pfam: PF01483; P\_1.  
CC Pfam: PF00082; peptidase\_S8; 1.  
CC PRINTS: PR00723; SUBTILISIN.  
CC ProDom: PD000717; P\_domain; 1.  
CC SMART: SM00181; EGF; 3.  
CC SMART: SM00001; EGF-like; 2.  
CC SMART: SM00261; FU; 22.  
CC PROSITE: PS00136; SUBTILASE\_ASP; 1.  
CC PROSITE: PS00137; SUBTILASE\_HIS; 1.  
CC PROSITE: PS00138; SUBTILASE\_SER; 1.  
CC Hydrolase: Serine protease; Glycoprotein; zymogen; signal;  
CC Cleavage on pair of basic residues; Repeat; Alternative splicing;  
CC Transmembrane.  
CC SIGNAL  
CC FT 1 34  
CC PROPEP 35 116  
CC CHAIN 117 1877

```

FT DOMAIN 117 1768 TYPE 5.
FT TRANSMEM 1769 1789 EXTRACELLULAR (POTENTIAL).
FT 1790 1877 POTENTIAL.
FT DOMAIN 117 452 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 464 602 CATALYTIC.
FT DOMAIN 638 1753 HOMO B.
FT DOMAIN 1825 1844 CYS-RICH MOTIF (CRM) REGION.
FT DOMAIN 1856 1877 AC 1.
FT SITE 116 117 AC 2.
FT SITE 521 523 CLEAVAGE (AUTO-) (BY SIMILARITY).
FT ACT_SITE 173 173 CELL ATTACHMENT SITE (POTENTIAL).
FT ACT_SITE 214 214 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT ACT_SITE 388 388 CHARGE RELAY SYSTEM (BY SIMILARITY).
FT CARBOHYD 227 227 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 383 383 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 667 667 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 804 804 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 854 854 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 951 951 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1016 1016 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1220 1220 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1317 1317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1523 1523 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1711 1711 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1733 1733 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT VARSPPLIC 878 915 GEYIDDOGHGRCFCEACSCAMGPTQDCISCPTRVLD ->
FT (IN ISOFORM PC5A).
FT VARSPPLIC 916 1877 MISSING (IN ISOFORM PC5A).
SQ SEQUENCE 1877 AA; 209287 MW; EC850E2DF20EALC3 CRC64;

```

```

Query Match 17.4%; Score 86.5; DB 1; Length 1877;
Best Local Similarity 22.1%; Pred. No. 0.36; Mismatches 43; Indels 53; Gaps 3;
Matches 29; Conservative 6;

```

```

OY 1 CSONEYFDSLHACIPQRLR--CSSNTPPLT---CORCEYFDSL----- 42
DB 1481 CAVEVWDESHRCOPCHKSCSGSEDCYCTCREFETLLMTGTCVKECPGHTDKDS 1540
OY 43 -----HACPCLRCSP-----PTCOYCFHSEYFDSLHACP----- 73
DB 1541 QOCVLCHSQRCECPHSMQCLSCRPGMFQLQCKHDYGGESTSGRCEKDKSK 1600
OY 74 -----PATCO 78
DB 1601 SCRGPRPTDQ 1611

RESULT 6
KRA3_SHEEP STANDARD; PRT; 131 AA.
AC P02441;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-JUL-1986 (Rel. 01, Last sequence update)
DT 01-AUG-1991 (Rel. 19, Last annotation update)
DE Keratin, high-sulfur matrix protein, IILN3.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE.
RA MEDLINE=74022242; PubMed=4584026;
RA Swart L.S., Haylett T.;
RT *Studies on the high-sulphur proteins of reduced Merino wool. Amino
RT acid sequence of protein SCMB-3A3."
RL Blochm J. 133:641-654(1973).
CC -I- FUNCTION: THE KERATIN PRODUCTS OF MAMMALIAN EPIDERMAL DERIVATIVES
CC SUCH AS WOOL AND HAIR CONSIST OF MICROFIBRILS EMBEDDED IN A RIGID

```

```

CC MATRIX OF OTHER PROTEINS. THE MATRIX PROTEINS INCLUDE THE HIGH-
CC SULFUR AND HIGH-TYROSINE KERATINS, HAVING MOLECULAR WEIGHTS OF
CC 6-20 kDa, WHEREAS THE MICROFIBRILS CONTAIN THE LARGER, LOW-SULFUR
CC KERATINS (40-56 kDa).
CC -I- MISCELLANEOUS: THE SOURCE OF THIS PROTEIN IS MERINO WOOL.
DR PIR: A02840; KRSHA3.
DR InterPro: IPR002494; Keratin_B2.
DR Pfam: PF01500; Keratin_B2.1.
KW Keratin.
SQ SEQUENCE 131 AA; 14163 MW; 8C0B56C0070B1C50 CRC64;

```

```

Query Match 16.7%; Score 83; DB 1; Length 131;
Best Local Similarity 25.2%; Pred. No. 0.082;
Matches 26; Conservative 12; Mismatches 35; Indels 30; Gaps 6;

```

```

OY 1 CSONEYFDSLHACIPQRLR--CSSNTPPLT-----CQR-YCCEYFDSLHA 44
DB 18 CLQPRVYRD-----PCCRPVSCQTVSRPVTVPCTRPIRCERCRPVCCPSIQEGC 71
OY 45 CPCLRCSPPCQY-----CCFHSEYFDSLHACP---PATCOP 79
DB 72 CPPTCCPTSCQAVVCRPCCATTCQCPVSVQCPCRCPTSCOP 114

```

```

RESULT 7
UBPM_MOUSE STANDARD; PRT; 526 AA.
AC 061068;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Ubiquitin carboxyl-terminal hydrolase DUB-1 (EC 3.1.2.15) (Ubiquitin
DE thiolesterase DUB-1) (Ubiquitin-specific processing protease DUB-1)
DE (Deubiquitinating enzyme 1).
GN DUB1 OR DUB-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96194957; PubMed=8622927;
RA Zhu Y., Carroll M., Papa F.R., Hochstrasser M., D'Andrea A.D.;
RT "DUB-1, a deubiquitinating enzyme with growth-suppressing activity."
RL Proc. Natl. Acad. Sci. U.S.A. 93:3275-3279(1996).
CC -I- FUNCTION: HAS GROWTH-SUPPRESSING ACTIVITY, INDUCES ARREST IN G1
CC PHASE UPON CONTROLLED EXPRESSION.
CC -I- CATALYTIC ACTIVITY: Ubiquitin C-terminal thiolester + H(2)O =
CC ubiquitin + a thiol.
CC -I- INDUCTION: BY INTERLEUKIN-3.
CC -I- SIMILARITY: BELONGS TO PEPTIDASE FAMILY C19. ALSO KNOWN AS
CC FAMILY 2 OF UBIQUITIN CARBOXYL-TERMINAL HYDROLASES.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U41636; AAC52532.1; -
CC MEROPS: C19.031; -
DR MGD; MGI:107699; DUB1.
DR InterPro: IPR001394; UCH-2.
DR Pfam: PF00443; UCH-1; 1.
DR PROSITE: PS00972; UCH_2_1; 1.
DR PROSITE: PS00973; UCH_2_2; 1.
DR PROSITE: PS0235; UCH_2_3; 1.
KW Ubiquitin conjugation; Hydrolase; Thiol protease; Multiene family.
FT ACT_SITE 60 60 BY SIMILARITY.

```

FT ACT\_SITE 298 298 BY SIMILARITY.  
DR ACT\_SITE 307 307 BY SIMILARITY.  
SQ SEQUENCE 526 AA; 59073 MW; 263AA7B757964FEA CRC64;  
Query Match 16.5%; Score 82; DB 1; Length 526;  
Best Local Similarity 35.8%; Pred. No. 0.33;  
Matches 19; Conservative 8; Mismatches 18; Indels 8; Gaps 2;  
OY 19 LRCSSNPPLTCORCYCEYFDSLHACPLRCSPPTCCYCFHSEYFDSLHA 71  
DB 66 LQCLHHTPL-----ADWMLSGHSGTC--CSPECKCKLAMEALVTQSLHS 110  
RESULT 8  
LMG3\_HUMAN STANDARD; PRT; 1587 AA.  
AC O9Y6N6:  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Laminin gamma-3 chain precursor (Laminin 12 gamma 3).  
GN LAMC3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Placenta;  
RX MEDLINE=99242614; PubMed=10225960;  
RA Koch M., Olson P., Albus A., Jin W., Hunter D.D., Brunken W.J.,  
RA Burgess R.E., Champiaud M.F.;  
RT "Characterization and expression of the laminin gamma3 chain: a novel,  
RT non-basement membrane-associated, laminin chain.";  
RL J. Cell Biol. 145:605-618(1999).  
CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR. LAMININ  
CC IS THOUGHT TO MEDiate THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
CC THE GAMMA-3 CHAIN IS A SUBUNIT OF LAMININ-12.  
CC -1- SUBCELLULAR LOCATION: Extracellular.  
CC -1- TISSUE SPECIFICITY: Broadly expressed in: skin, heart, lung, and  
CC the reproductive tracts.  
CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
CC -1- DOMAIN: DOMAIN IV IS GLOBULAR.  
CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
CC -1- SIMILARITY: CONTAINS 1 LAMININ EGF-LIKE DOMAINS.  
CC -1- SIMILARITY: CONTAINS 1 LAMININ DOMAIN IV.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL: AF041835; AAD36991.1; -  
DR HSSP: P02468; 1TLE.  
DR MIM: 604349; -  
DR InterPro: IPR0000561; EGF-like.  
DR InterPro: IPR000034; Laminin\_B.  
DR InterPro: IPR002049; Laminin\_EGF.  
DR InterPro: IPR001886; LAMNT.  
DR Pfam: PF00052; laminin\_B\_1.  
DR Pfam: PF00053; laminin\_EGF\_10.  
DR Pfam: PF00055; laminin\_Nterm; 1.

DR PRINTS: PR00011; EGF/LAMININ.  
DR ProDom: PD002082; LAMNT; 1.  
DR ProDom: PD003031; Laminin\_B; 1.  
DR SMART: SM00180; EGF\_Lam; 10.  
DR SMART: SM00001; EGF\_Like; 1.  
DR SMART: SM00281; Lam; 1.  
DR SMART: SM00136; LAMNT; 1.  
DR PROSITE: PS00022; EGF\_1; 7.  
DR PROSITE: PS01186; EGF\_2; 2.  
DR PROSITE: PS01248; LAMININ\_Type\_EGF; 10.  
KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
FT SIGNAL 1 19  
FT CHAIN 1 19  
FT DOMAIN 20 1587  
FT DOMAIN 20 270  
FT DOMAIN 271 326  
FT DOMAIN 327 382  
FT DOMAIN 383 429  
FT DOMAIN 430 479  
FT DOMAIN 480 489  
FT DOMAIN 490 672  
FT DOMAIN 673 706  
FT DOMAIN 707 754  
FT DOMAIN 755 809  
FT DOMAIN 810 865  
FT DOMAIN 866 916  
FT DOMAIN 917 964  
FT DOMAIN 965 1013  
FT DOMAIN 1014 1587  
FT DOMAIN 1071 1141  
FT DOMAIN 1200 1229  
FT DOMAIN 1424 1504  
FT DOMAIN 1535 1579  
FT SITE 1059 1061  
FT CARBOHD 87 87  
FT CARBOHD 119 119  
FT CARBOHD 295 295  
FT CARBOHD 328 328  
FT CARBOHD 631 631  
FT CARBOHD 837 837  
FT CARBOHD 980 980  
FT CARBOHD 1185 1185  
FT CARBOHD 1518 1518  
SQ SEQUENCE 1587 AA; 172051 MW; 3CB6E0B5F20319 CRC64;  
Query Match 16.3%; Score 81; DB 1; Length 1587;  
Best Local Similarity 31.0%; Pred. No. 1.1;  
Matches 22; Conservative 8; Mismatches 31; Indels 10; Gaps 4;  
OY 12 HACIPQQLRCSSNPPLTCORCYCEYFDSLHACPLRC---SPPTCCYCFHSEYFDS 67  
DB 322 HECLPC--NCSGRSECFDR---ELFRSTGSHGRCHRCRDHTAGPHRCQENFYHWP 376  
OY 68 LHACRPATCQ 78  
DB 377 RM-PCQPCDQ 386  
RESULT 9  
PAC4\_RAT STANDARD; PRT; 937 AA.  
ID PAC4\_RAT  
AC Q63415;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Paired basic amino acid cleaving enzyme 4 precursor (EC 3.4.21.-)  
DE (Subtilisin/kexin-like protease PAC4) (Subtilisin-like proprotein  
DE convertase 4) (SPC4).  
DE PACE4.  
GN PACE4.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Hypothalamus, and Pituitary;  
 RX MEDLINE=94349873; PubMed=8070361;  
 RA Johnson R.C., Darlington D.N., Hand T.A., Bloomquist B.T., Mains R.E.;  
 RT "PAC4: a subtilisin-like endoprotease prevalent in the anterior  
 pituitary and regulated by thyroid status.";  
 RL Endocrinology 135:1178-1185(1994).  
 CC -1- FUNCTION: LIKELY TO REPRESENT AN ENDOPROTEASE WITHIN THE  
 CC CONSTITUTIVE SECRETORY PATHWAY WITH UNIQUE RESTRICTED  
 CC DISTRIBUTION IN BOTH NEUROENDOCRINE AND NON-NEUROENDOCRINE TISSUES  
 CC AND CAPABLE OF CLEAVAGE AT THE RX(R)/R/R CONSENSUS MOTIF.  
 CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR  
 CC PROTEIN BY CLEAVAGE OF ARG-XAA-YAA-ARG-1-ZAA BONDS,  
 CC WHERE XAA CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.  
 CC -1- COFACTOR: PROBABLY CALCIUM-DEPENDENT (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN THE ANTERIOR PITUITARY AND  
 CC IN SEVERAL BRAIN REGIONS, THE ATRIUM, AND THE VENTRICLE.  
 CC -1- DOMAIN: THE PROPEPTIDE DOMAIN ACTS AS AN INTRAMOLECULAR CHAPERONE  
 CC ASSISTING THE FOLDING OF THE ZMOGEN WITHIN THE ENDOPLASMIC  
 CC RETICULUM.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S8; ALSO KNOWN AS THE  
 CC SUBTILISIN.  
 CC -1- SIMILARITY: CONTAINS 1 HOMO B/P DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: L31894; AA61987.1; -  
 DR HSP: 099405; 1MPt.  
 DR MEROPS: S08.075; -  
 DR InterPro: IPR000661; EGF-like.  
 DR InterPro: IPR002174; Furin-like.  
 DR InterPro: IPR002884; P\_domain.  
 DR InterPro: IPR000209; Peptidase\_S8.  
 DR Pfam: PF01483; P\_1.  
 DR PRINTS: PR00082; Peptidase\_S8; 1.  
 DR PRINTS: PR00723; SUBTILISIN.  
 DR ProDom: PD000717; P\_domain; 1.  
 DR SMART: SM00181; EGF; 1.  
 DR SMART: SM00261; FU; 5.  
 DR PROSITE: PS00136; SUBTILISIN-ASP; 1.  
 DR PROSITE: PS00137; SUBTILISIN-HIS; 1.  
 DR PROSITE: PS00138; SUBTILISIN-SER; 1.  
 KW Hydrolyase; Serine protease; Glycoprotein; Zymogen; Signal; Calcium;  
 KW Cleavage on pair of basic residues; Repeat.  
 FT SIGNAL 1  
 FT PROPEP 45  
 FT CHAIN 133  
 FT DOMAIN 133  
 FT DOMAIN 133  
 FT SITE 133  
 FT SITE 133  
 FT SITE 536  
 FT ACT\_SITE 186  
 FT ACT\_SITE 186  
 FT ACT\_SITE 227  
 FT ACT\_SITE 227  
 FT ACT\_SITE 401  
 FT ACT\_SITE 401  
 FT CARBOHYD 240  
 FT CARBOHYD 240  
 FT CARBOHYD 882  
 FT CARBOHYD 882  
 FT CARBOHYD 900  
 FT CARBOHYD 900  
 SQ SEQUENCE 937 AA; 104053 MW; F3865557C3705C8 CRC64;

QY 1 CSQNEPDSLHACIPOLRCS-----SNTPLTCQR--YCSEYFDSLHACP----- 46  
 Db 701 CPGLYFEDTARRRORRHKHKGCEPTGTGSPGOCSCSRGFGFHHQETNVCVLLCPAGLAD 760  
 QY 47 ---CLRCSPPTCYC-----CFHSEYDSLHACPPATQC 78  
 Db 761 SQRLCLRHG-PSQCKVDEPEKSTVCKEFGSLARGSCIPDCEPTGYDSEILRC--GECH 817  
 QY 79 PYC 81  
 Db 818 HTC 820  
 RESULT 10  
 LMB2\_HUMAN STANDARD; PRT; 1798 AA.  
 ID P5268; Q16321;  
 AC 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Laminin beta-2 chain precursor (S-laminin) (Laminin B1s chain).  
 GN LAMB2 OR LAMS.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95213013; PubMed=7698745;  
 RA Wewer U.M., Gerecke D.R., Durkin M.E., Kurtz K.S., Mattei M.-G.,  
 RA Champland M.F., Burgeson R.E., Albrechtsen R.;  
 RT "Human beta 2 chain of laminin (formerly S chain): cDNA cloning,  
 RT chromosomal localization, and expression in carcinomas.";  
 RL Genomics 24:243-252(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95316263; PubMed=7795887;  
 RA Ilvainen A., Vuolteenaho R., Sainio K., Eddy R., Shows T.B.,  
 RA Sariola H., Tryggvason K.;  
 RT "The human laminin beta 2 chain (S-laminin): structure, expression in  
 RT fetal tissues and chromosomal assignment of the LAMB2 gene.";  
 RL Matrix Biol. 14:489-497(1995).  
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
 CC THE BETA-2 CHAIN IS A SUBUNIT OF LAMININ-3 (S-LAMININ), LAMININ-4  
 CC (S-MEROSIN), AND LAMININ-7 (KS-LAMININ).  
 CC -1- SUBCELLULAR LOCATION: Extracellular.  
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR  
 CC COMPONENT). S-LAMININ IS CONCENTRATED IN THE SYNAPTIC  
 CC CLEFT OF THE NEUROMUSCULAR JUNCTION.  
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
 CC -1- SIMILARITY: CONTAINS 12.5 LAMININ EGF-LIKE DOMAINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z68155; CAA92279.1; -

Query Match 16.2%; Score 80.5; DB 1; Length 937;  
 Best Local Similarity 24.4%; Pred. No. 0.75; Mismatches 39; Indels 45; Gaps 6;  
 Matches 30; Conservative 9;

DR EMBL: 268156; CA92279.1; JOINED.  
 DR EMBL: X79683; CA56130.1; -.  
 DR EMBL: S7512; AAB34682.2; -.  
 DR HSSP: P02468; IKLO.  
 DR MIM: 150325; -.  
 DR InterPro: IPR000561; EGF-Like.  
 DR InterPro: IPR00186; LamNT.  
 DR InterPro: IPR02049; Laminin\_EGF.  
 DR Pfam: PF00053; Laminin\_EGF\_13.  
 DR Pfam: PF00055; Laminin\_Nterm; 1.  
 DR PRINTS: PRO0011; EGFLAMININ.  
 DR ProDom: PD002082; LamNT; 1.  
 DR SMART: SM00180; EGF\_Lam; 11.  
 DR SMART: SM00136; LamNT; 1.  
 DR PROSITE: PS00022; EGF\_1; 10.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01248; LAMININ TYPE EGF; 12.  
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 FT SIGNAL 1 32  
 FT CHAIN 33 1798  
 FT DOMAIN 33 280 LAMININ BETA-2 CHAIN.  
 FT DOMAIN 283 346 LAMININ N-TERMINAL (DOMAIN VI).  
 FT DOMAIN 347 409 LAMININ EGF-LIKE 1.  
 FT DOMAIN 410 469 LAMININ EGF-LIKE 2.  
 FT DOMAIN 470 521 LAMININ EGF-LIKE 3.  
 FT DOMAIN 522 552 LAMININ EGF-LIKE 4.  
 FT DOMAIN 553 781 LAMININ EGF-LIKE 5 (INCOMPLETE).  
 FT DOMAIN 783 830 LAMININ DOMAIN IV.  
 FT DOMAIN 831 876 LAMININ EGF-LIKE 6.  
 FT DOMAIN 877 926 LAMININ EGF-LIKE 7.  
 FT DOMAIN 927 985 LAMININ EGF-LIKE 8.  
 FT DOMAIN 986 1037 LAMININ EGF-LIKE 9.  
 FT DOMAIN 1038 1094 LAMININ EGF-LIKE 10.  
 FT DOMAIN 1095 1142 LAMININ EGF-LIKE 11.  
 FT DOMAIN 1143 1189 LAMININ EGF-LIKE 12.  
 FT DOMAIN 1190 1409 LAMININ EGF-LIKE 13.  
 FT DOMAIN 1410 1442 DOMAIN II.  
 FT DOMAIN 1443 1798 DOMAIN ALPHA.  
 FT DOMAIN 1253 1319 COILED COIL (POTENTIAL).  
 FT DOMAIN 1472 1526 COILED COIL (POTENTIAL).  
 FT DOMAIN 1577 1790 COILED COIL (POTENTIAL).  
 FT DISULFD 283 292 BY SIMILARITY.  
 FT DISULFD 285 310 BY SIMILARITY.  
 FT DISULFD 312 321 BY SIMILARITY.  
 FT DISULFD 324 344 BY SIMILARITY.  
 FT DISULFD 347 356 BY SIMILARITY.  
 FT DISULFD 349 374 BY SIMILARITY.  
 FT DISULFD 377 386 BY SIMILARITY.  
 FT DISULFD 389 407 BY SIMILARITY.  
 FT DISULFD 410 423 BY SIMILARITY.  
 FT DISULFD 412 438 BY SIMILARITY.  
 FT DISULFD 440 449 BY SIMILARITY.  
 FT DISULFD 452 467 BY SIMILARITY.  
 FT DISULFD 470 484 BY SIMILARITY.  
 FT DISULFD 472 491 BY SIMILARITY.  
 FT DISULFD 493 502 BY SIMILARITY.  
 FT DISULFD 505 519 BY SIMILARITY.  
 FT DISULFD 783 795 BY SIMILARITY.  
 FT DISULFD 785 802 BY SIMILARITY.  
 FT DISULFD 804 813 BY SIMILARITY.  
 FT DISULFD 816 828 BY SIMILARITY.  
 FT DISULFD 831 843 BY SIMILARITY.  
 FT DISULFD 833 850 BY SIMILARITY.  
 FT DISULFD 852 861 BY SIMILARITY.  
 FT DISULFD 864 874 BY SIMILARITY.  
 FT DISULFD 877 886 BY SIMILARITY.  
 FT DISULFD 879 893 BY SIMILARITY.  
 FT DISULFD 896 905 BY SIMILARITY.  
 FT DISULFD 908 924 BY SIMILARITY.  
 FT DISULFD 927 943 BY SIMILARITY.  
 FT DISULFD 929 954 BY SIMILARITY.  
 FT DISULFD 956 965 BY SIMILARITY.

FT DISULFD 968 983 BY SIMILARITY.  
 FT DISULFD 986 1000 BY SIMILARITY.  
 FT DISULFD 988 1007 BY SIMILARITY.  
 FT DISULFD 1010 1019 BY SIMILARITY.  
 FT DISULFD 1022 1035 BY SIMILARITY.  
 FT DISULFD 1095 1107 BY SIMILARITY.  
 FT DISULFD 1097 1114 BY SIMILARITY.  
 FT DISULFD 1116 1125 BY SIMILARITY.  
 FT DISULFD 1128 1140 BY SIMILARITY.  
 FT DISULFD 1143 1155 BY SIMILARITY.  
 FT DISULFD 1145 1162 BY SIMILARITY.  
 FT DISULFD 1164 1173 BY SIMILARITY.  
 FT DISULFD 1176 1187 BY SIMILARITY.  
 FT DISULFD 1190 1190 INTERCHAIN (PROBABLE).  
 FT DISULFD 1193 1193 INTERCHAIN (PROBABLE).  
 FT DISULFD 1797 1797 INTERCHAIN (PROBABLE).  
 FT CARBOHYD 248 248 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 368 368 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1085 1085 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1249 1249 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1308 1308 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1348 1348 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 1499 1499 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CONFLICT 914 914 R -> G (IN REF. 2).  
 FT CONFLICT 1179 1179 G -> A (IN REF. 2).  
 SQ SEQUENCE 1798 AA; 196079 MW; 9355CE5B24850CB7 CRC64;  
 Query Match 16.2%; Score 80.5; DB 1; Length 1798;  
 Best local similarity 26.9%; Pred. No. 1.3;  
 Matches 29; Conservative 7; Mismatches 31; Indels 41; Gaps 8;  
 QY 1 CSQNEYFDSL-----HACIP-----COL-RCSSNTPPLTCQRYCC 35  
 DB 943 CHQDEYSQIYCHCRAGYGLRCEACAPGHGDFSPGRCQLCECSGNDPMDPD--AC 1000  
 QY 36 EYFDSLHACPLRC-----SPTCOYC--CFHSEYFDSLHACPRATC 77  
 DB 1001 D-----PHTGQCLRLHTEGPHCAKCRFGHQAAROSCHRC--TC 1040  
 RESULT 11  
 LMA1\_HUMAN STANDARD: PRT; 3075 AA.  
 AC P25391;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Laminin alpha-1 chain precursor (Laminin A chain).  
 GN LAMA1 OR LAMA.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91333420; PubMed=1714537;  
 RA Haaparanta T., Uitto J., Ruoslahti E., Engvall E.;  
 RT "Molecular cloning of the cDNA encoding human laminin A chain.";  
 RL Matrix 11:151-160(1991).  
 RM [2]  
 RP SEQUENCE OF 1-2628 FROM N.A.  
 RX MEDLINE=91264789; PubMed=2049067;  
 RA Nissinen M., Vuolteenaho R., Booc-Handford R., Kallunki P.,  
 RA Tryggvason K.;  
 RT "Primary structure of the human laminin A chain. Limited expression  
 in human tissues.";  
 RL Biochem. J. 276:369-379(1991).  
 RM [3]  
 RP SEQUENCE OF 2397-3072 FROM N.A.  
 RX MEDLINE=89280632; PubMed=2733383;  
 RA Olsen D., Nagayoshi T., Fazio M., Peltonen J., Jaakkola S.,  
 RA Sandorn D., Sasaki T., Kuivanen H., Chn M.L., Deutzmann R.,

RA Timpl R., Utto B.:  
 RT "Human laminin: Cloning and sequence analysis of cDNAs encoding A, B1  
 RT and B2 chains, and expression of the corresponding genes in human  
 RT skin and cultured cells."  
 RT Lab. Invest. 60:772-782(1989).  
 CC -1- FUNCTION: BINDING TO CELLS VIA A HIGH AFFINITY RECEPTOR, LAMININ  
 CC IS THOUGHT TO MEDIATE THE ATTACHMENT, MIGRATION, & ORGANIZATION OF  
 CC CELLS INTO TISSUES DURING EMBRYONIC DEVELOPMENT BY INTERACTING  
 CC WITH OTHER EXTRACELLULAR MATRIX COMPONENTS.  
 CC -1- SUBUNIT: LAMININ IS A COMPLEX GLYCOPROTEIN, CONSISTING OF THREE  
 CC DIFFERENT POLYPEPTIDE CHAINS (ALPHA, BETA, GAMMA), WHICH ARE BOUND  
 CC TO EACH OTHER BY DISULFIDE BONDS INTO A CROSS-SHAPED MOLECULE  
 CC COMPRISING ONE LONG & THREE SHORT ARMS WITH GLOBULES AT EACH END.  
 CC THE ALPHA-1 CHAIN IS A SUBUNIT OF LAMININ-1 (EHS LAMININ) AND  
 CC LAMININ-3 (SL LAMININ).  
 CC -1- TISSUE SPECIFICITY: Extracellular.  
 CC -1- TISSUE SPECIFICITY: FOUND IN THE BASEMENT MEMBRANES (MAJOR  
 CC COMPONENT).  
 CC -1- DOMAIN: THE ALPHA-HELICAL DOMAINS I AND II ARE THOUGHT TO INTERACT  
 CC WITH OTHER LAMININ CHAINS TO FORM A COILED COIL STRUCTURE.  
 CC -1- SIMILARITY: DOMAINS VI, IV AND G ARE GLOBULAR.  
 CC -1- SIMILARITY: CONTAINS 1 LAMININ N-TERMINAL DOMAIN (DOMAIN VI).  
 CC -1- SIMILARITY: CONTAINS 17 LAMININ EGF-LIKE DOMAINS.  
 CC -1- SIMILARITY: CONTAINS 5 LAMININ G-LIKE DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X58531; CAA61418.1; -  
 DR PIR: S14458; S14458.  
 DR HSSP: P02468; IKID.  
 DR MIM: 150320; -  
 DR InterPro: IPR000561; EGF-like.  
 DR InterPro: IPR001886; LAMNT.  
 DR InterPro: IPR000084; Laminin\_B.  
 DR InterPro: IPR002069; Laminin\_EGF.  
 DR InterPro: IPR001791; Laminin\_B; 2.  
 DR Pfam: PF00052; Laminin\_B; 2.  
 DR Pfam: PF00053; Laminin\_EGF; 15.  
 DR Pfam: PF00054; Laminin\_G; 5.  
 DR Pfam: PF00055; Laminin\_Nterm; 1.  
 DR PRINTS: PRO0011; EGF\_LAMININ.  
 DR PRODom: PD002082; LAMNT; 1.  
 DR PRODom: PD003031; Laminin\_B; 2.  
 DR SMART: SM00180; EGF\_Lam; 14.  
 DR SMART: SM00001; EGF\_Like; 1.  
 DR SMART: SM00281; Lamb; 2.  
 DR SMART: SM00282; Lamb; 5.  
 DR SMART: SM00136; LAMNT; 1.  
 DR PROSITE: PS00022; EGF\_1; 11.  
 DR PROSITE: PS01186; EGF\_2; 2.  
 DR PROSITE: PS01248; LAMININ\_TYPE\_EGF; 15.  
 DR PROSITE: PS50025; LAM\_G\_DOMAIN; 5.  
 KW Glycoprotein; Basement membrane; Extracellular matrix; Coiled coil;  
 KW Laminin EGF-like domain; Cell adhesion; Repeat; Signal.  
 FT SIGNAL 1  
 FT CHAIN 17  
 FT 18 3075 LAMININ ALPHA-1 CHAIN.  
 FT 18 3075 LAMININ N-TERMINAL (DOMAIN VI).  
 FT 269 326 LAMININ EGF-LIKE 1.  
 FT 270 326 LAMININ EGF-LIKE 1.  
 FT 327 396 LAMININ EGF-LIKE 2.  
 FT 397 453 LAMININ EGF-LIKE 3.  
 FT 454 502 LAMININ EGF-LIKE 3.  
 FT 503 512 LAMININ EGF-LIKE 5 (N-TERMINAL).  
 FT 517 708 LAMININ DOMAIN IV 1 (DOMAIN IV B).  
 FT 709 741 LAMININ EGF-LIKE 5 (C-TERMINAL).  
 FT 742 790 LAMININ EGF-LIKE 6.  
 FT 791 848 LAMININ EGF-LIKE 7.  
 FT 848 LAMININ EGF-LIKE 7.

FT 849 901 LAMININ EGF-LIKE 8.  
 FT 902 950 LAMININ EGF-LIKE 9.  
 FT 951 997 LAMININ EGF-LIKE 10.  
 FT 998 1043 LAMININ EGF-LIKE 11.  
 FT 1044 1089 LAMININ EGF-LIKE 12.  
 FT 1090 1149 LAMININ EGF-LIKE 13.  
 FT 1150 1159 LAMININ EGF-LIKE 14 (N-TERMINAL).  
 FT 1160 1361 LAMININ DOMAIN IV 2 (DOMAIN IV A).  
 FT 1362 1402 LAMININ EGF-LIKE 14 (C-TERMINAL).  
 FT 1403 1451 LAMININ EGF-LIKE 15.  
 FT 1452 1508 LAMININ EGF-LIKE 16.  
 FT 1509 1556 LAMININ EGF-LIKE 17.  
 FT 1556 2116 DOMAIN II AND I.  
 FT 2117 2297 LAMININ G-LIKE 1.  
 FT 2297 2481 LAMININ G-LIKE 2.  
 FT 2481 2673 LAMININ G-LIKE 3.  
 FT 2673 2885 LAMININ G-LIKE 4.  
 FT 2885 3070 LAMININ G-LIKE 5.  
 FT 3070 3270 LAMININ G-LIKE 6.  
 FT 3270 3366 COILED COIL (POTENTIAL).  
 FT 3366 361 COILED COIL (POTENTIAL).  
 FT 361 373 COILED COIL (POTENTIAL).  
 FT 373 394 COILED COIL (POTENTIAL).  
 FT 394 409 CELL ATTACHMENT SITE.  
 FT 409 427 BY SIMILARITY.  
 FT 427 438 BY SIMILARITY.  
 FT 438 441 BY SIMILARITY.  
 FT 441 451 BY SIMILARITY.  
 FT 451 467 BY SIMILARITY.  
 FT 467 471 BY SIMILARITY.  
 FT 471 482 BY SIMILARITY.  
 FT 482 485 BY SIMILARITY.  
 FT 485 500 BY SIMILARITY.  
 FT 500 504 BY SIMILARITY.  
 FT 504 514 BY SIMILARITY.  
 FT 514 524 BY SIMILARITY.  
 FT 524 536 BY SIMILARITY.  
 FT 536 549 BY SIMILARITY.  
 FT 549 561 BY SIMILARITY.  
 FT 561 571 BY SIMILARITY.  
 FT 571 581 BY SIMILARITY.  
 FT 581 591 BY SIMILARITY.  
 FT 591 601 BY SIMILARITY.  
 FT 601 611 BY SIMILARITY.  
 FT 611 621 BY SIMILARITY.  
 FT 621 631 BY SIMILARITY.  
 FT 631 641 BY SIMILARITY.  
 FT 641 651 BY SIMILARITY.  
 FT 651 661 BY SIMILARITY.  
 FT 661 671 BY SIMILARITY.  
 FT 671 681 BY SIMILARITY.  
 FT 681 691 BY SIMILARITY.  
 FT 691 701 BY SIMILARITY.  
 FT 701 711 BY SIMILARITY.  
 FT 711 721 BY SIMILARITY.  
 FT 721 731 BY SIMILARITY.  
 FT 731 741 BY SIMILARITY.  
 FT 741 751 BY SIMILARITY.  
 FT 751 761 BY SIMILARITY.  
 FT 761 771 BY SIMILARITY.  
 FT 771 781 BY SIMILARITY.  
 FT 781 791 BY SIMILARITY.  
 FT 791 801 BY SIMILARITY.  
 FT 801 811 BY SIMILARITY.  
 FT 811 821 BY SIMILARITY.  
 FT 821 831 BY SIMILARITY.  
 FT 831 841 BY SIMILARITY.  
 FT 841 851 BY SIMILARITY.  
 FT 851 861 BY SIMILARITY.  
 FT 861 870 BY SIMILARITY.  
 FT 870 882 BY SIMILARITY.  
 FT 882 895 BY SIMILARITY.  
 FT 895 904 BY SIMILARITY.  
 FT 904 914 BY SIMILARITY.  
 FT 914 921 BY SIMILARITY.  
 FT 921 932 BY SIMILARITY.  
 FT 932 948 BY SIMILARITY.  
 FT 948 963 BY SIMILARITY.  
 FT 963 980 BY SIMILARITY.  
 FT 980 995 BY SIMILARITY.  
 FT 995 1007 BY SIMILARITY.  
 FT 1007 1014 BY SIMILARITY.  
 FT 1014 1025 BY SIMILARITY.  
 FT 1025 1041 BY SIMILARITY.  
 FT 1041 1056 BY SIMILARITY.  
 FT 1056 1063 BY SIMILARITY.  
 FT 1063 1074 BY SIMILARITY.  
 FT 1074 1087 BY SIMILARITY.  
 FT 1087 1107 BY SIMILARITY.  
 FT 1107 1142 BY SIMILARITY.  
 FT 1142 1191 BY SIMILARITY.  
 FT 1191 1222 BY SIMILARITY.  
 FT 1222 1249 BY SIMILARITY.  
 FT 1249 1266 BY SIMILARITY.  
 FT 1266 1283 BY SIMILARITY.  
 FT 1283 1301 BY SIMILARITY.  
 FT 1301 1318 BY SIMILARITY.  
 FT 1318 1335 BY SIMILARITY.  
 FT 1335 1352 BY SIMILARITY.  
 FT 1352 1369 BY SIMILARITY.  
 FT 1369 1386 BY SIMILARITY.  
 FT 1386 1403 BY SIMILARITY.  
 FT 1403 1420 BY SIMILARITY.  
 FT 1420 1437 BY SIMILARITY.  
 FT 1437 1454 BY SIMILARITY.  
 FT 1454 1476 BY SIMILARITY.

FT DISULFID 1479 1488 BY SIMILARITY.  
FT DISULFID 1491 1506 BY SIMILARITY.  
FT DISULFID 1509 1521 BY SIMILARITY.  
FT DISULFID 1511 1528 BY SIMILARITY.  
FT DISULFID 1530 1539 BY SIMILARITY.  
FT DISULFID 1542 1553 BY SIMILARITY.  
FT DISULFID 1556 1556 INTERCHAIN (PROBABLE).  
FT DISULFID 1560 1560 INTERCHAIN (PROBABLE).  
FT CARBOHYD 38 38 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 555 555 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 665 665 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 763 763 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 926 926 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 952 952 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1045 1045 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1407 1407 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1579 1579 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1596 1596 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1678 1678 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1689 1689 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1698 1698 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1717 1717 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1804 1804 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1894 1894 N-LINKED (GLCNAC. . .) (POTENTIAL).

Query Match 16.1%; Score 80; DB 1; Length 3075;  
Best Local Similarity 23.4%; Pred. No. 2.3;  
Matches 29; Conservative 10; Mismatches 37; Indels 48; Gaps 8;

OY 1 CSQNFPSL--HACIPCQ-----LRCSNTPPLTCQRYCCE-- 36  
DB 831 CADGYGNPTVGESECVPCDSCGNVDPSEAGCHDSVTGECCLKCLANTGACRER-CADGF 889  
OY 37 YFDSL--LHACPC-----LRCSNTPPLTCQRYCCE-- 73  
DB 890 YGDAVTAKNRACRCECHVKGSHSVCHELTGLCDCKRPNTGQCCDQCLHGVCYGLDSHGCR 949  
OY 74 PATC 77  
DB 950 PCNC 953

RESULT 12  
YNC3\_YEAST STANDARD; PRT; 965 AA.  
AC P53971;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Hypothetical 108.5 kDa protein in UME3-HDA1 intergenic region.  
GN YNL022C OR N2812.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
OX NCBI\_TaxID=4932;  
[1]  
RP SEQUENCE FROM N.A.  
RA Duestenheft A., Floeth M., Filtz C., Heuss-Neitzel D.,  
RA Hilbert H., Moestl D.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Andre B., Iratqui Housaini I., Urrestarazu L.A., Vissers S.;  
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: TO DROSOPHILA SHUTTLE CRAFT PROTEIN (STC) AND HUMAN  
CC NFX1.  
CC -1- SIMILARITY: CONTAINS 1 RING-TYPE ZINC FINGER.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: 271299; CAB95885.1; -  
DR SGD: S0004968; YNL023C.  
DR InterPro: IPR001374; R3H.  
DR InterPro: IPR000967; ZNF\_NFX1.  
DR InterPro: IPR001841; ZNF\_fing.  
DR Pfam: PF01424; R3H; 1.  
DR Pfam: PF01422; ZF-NFX-1; 5.  
DR SMART: SM00393; R3H; 1.  
DR SMART: SM00184; RING; 1.  
DR SMART: SM00438; ZNF\_NFX; 7.  
DR PROSITE: PS00518; ZF\_RING\_1; FALSE\_NEG.  
DR PROSITE: PS00889; ZF\_RING\_2; 1.  
KW Hypothetical protein; zinc-finger; Repeat.  
FT ZN\_FING 68 118 RING-TYPE  
FT DOMAIN 150 644 7 X APPROXIMATE REPEATS, CYS-RICH.  
FT REPEAT 150 185 1.  
FT REPEAT 206 251 2.  
FT REPEAT 273 330 3.  
FT REPEAT 352 390 4.  
FT REPEAT 458 497 5.  
FT REPEAT 575 610 6.  
FT REPEAT 611 644 7.  
SO SEQUENCE 965 AA; 108494 MW; 121C57BB07C6FA9D CRC64;

Query Match 16.0%; Score 79.5; DB 1; Length 965;  
Best Local Similarity 29.6%; Pred. No. 0.96;  
Matches 29; Conservative 6; Mismatches 26; Indels 37; Gaps 7;

OY 12 HACT-PCQLRCSSNTPPLTCQRYC-----CEYFDSLHACPC-----LRCSNTP 53  
DB 456 HICLPKMLTSLSCGII--KQARKCPKCPCLSDSDVLCPCGNVVPAPVRCGKLP 513  
OY 54 TCQYCCF-----HSEYFDSLHACPC 76  
DB 514 TCHPCTKVRGSESTCGHKPMPTCH--SLDVSCPCT 549

RESULT 13  
HMCT\_BOMMO STANDARD; PRT; 3133 AA.  
AC P96092;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Hemocytin precursor (Humoral lectin).  
OS Bombyx mori (Silk moth).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;  
OC Bombycoidea; Bombycidae; Bombyx.  
OX NCBI\_TaxID=7091;  
[1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-FUYOU X TOKAI; TISSUE-Hemocyte;  
RX MEDLINE=95178544; PubMed=7873598;  
RA Kotani E., Yamakawa M., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,  
RA Matsubara F., Tanai K., Kado-no-Okuda K., Kato Y., Mori H.;  
RT "Cloning and expression of the gene of hemocytin, an insect humoral  
RT lectin which is homologous with the mammalian von Willebrand  
RT factor";  
RL Blochim. Biophys. Acta 1260:245-258(1995).  
RN [2]  
RP SEQUENCE OF 2221-3133 FROM N.A.  
RA Kotani E., Iwamoto S.-I., Tashiro M., Mori H., Sumida M.,  
RA Matsubara F., Yamakawa M.;  
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: ADHESIVE PROTEIN AND RELATES TO HEMOSTASIS OR  
CC ENCASCULATION OF FOREIGN SUBSTANCES FOR SELF-DEFENSE.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED IN HEMOCYTES DURING LARVAL-PUPAL  
CC METAMORPHOSIS.

CC -1- INDUCTION: HEMAGGLUTINATION ACTIVITY IS INCREASED BY BACTERIAL,  
CC OR VIRAL INFECTION AND INHIBITED BY D-MANNOSE, N-ACETYL-D-  
CC GALACTOSAMINE AND D-MALTOSE.  
CC -1- PTH: MAY BE CONVERTED INTO THE 260 KDa MATURE HEMOCYTIN BY  
CC PROTEOLYSIS.  
CC -1- SIMILARITY: TO MAMMALIAN VON WILLEBRAND FACTOR.  
CC -1- SIMILARITY: THE N-TERMINAL DOMAIN SHOWS SOME SIMILARITY TO THAT  
CC OF HUMAN MUCIN 2.  
CC -1- SIMILARITY: CONTAINS 2 F5/8 TYPE C DOMAINS.  
CC -1- SIMILARITY: CONTAINS 1 C-TERMINAL CYSTINE KNOT-LIKE DOMAIN (CTCK).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: D29738; BABD6160.1; -  
CC EMBL: D14035; BABD3124.1; -  
CC InterPro: IPR000359; Cys\_knot.  
CC InterPro: IPR000421; FA58\_C.  
CC InterPro: IPR002172; LDL\_recept\_A.  
CC InterPro: IPR002919; TIL.  
CC InterPro: IPR001007; VMFC.  
CC InterPro: IPR001846; Vwd.  
CC InterPro: IPR001878; ZnF\_CCHC.  
CC InterPro: IPR000083; fibronectin\_type\_1.  
CC Pfam: PF00754; F5\_LF8\_type\_C; 2.  
CC Pfam: PF01826; TIL; 6.  
CC SMART: SM00094; vwd; 3.  
DR SMART: SM00041; CT; 1.  
DR SMART: SM00038; FN1; 1.  
DR SMART: SM00192; LDL; 1.  
DR SMART: SM00214; VWC; 3.  
DR SMART: SM00216; VWD; 3.  
DR SMART: SM00343; ZNF\_C2HC; 1.  
DR PROSITE: PS01185; CTCK\_1; 1.  
DR PROSITE: PS01208; VMFC; FALSE-NEG.  
DR PROSITE: PS01225; CTCK\_2; 1.  
DR PROSITE: PS01285; FA58C\_1; 2.  
DR PROSITE: PS01286; FA58C\_2; 2.  
KW Lectin; Glycoprotein; Signal; Repeat; Cell adhesion.  
FT SIGNAL 1  
FT CHAIN ?  
FT DOMAIN 29 3133  
FT DOMAIN 153 131  
FT DOMAIN 248 613  
FT DOMAIN 940 1095  
FT DOMAIN 1116 1254  
FT DOMAIN 1283 1386  
FT DOMAIN 1620 1951  
FT DOMAIN 1952 2315  
FT DOMAIN 2230 2321  
FT DOMAIN 2335 2361  
FT DOMAIN 2435 2469  
FT DOMAIN 2553 2622  
FT DOMAIN 2842 2907  
FT DOMAIN 2971 3076  
FT DOMAIN 895 914  
FT DOMAIN 1267 1270  
FT DOMAIN 1425 1428  
FT DOMAIN 1447 1450  
FT DOMAIN 1474 1479  
FT DOMAIN 2148 2153  
FT DOMAIN 2156 2159  
FT DOMAIN 2341 2344  
FT DISULFID 940 1095  
FT DISULFID 1116 1254  
FT DISULFID 2981 3040

FT DISULFID 2991 3054 BY SIMILARITY.  
FT DISULFID 3004 3070 BY SIMILARITY.  
FT DISULFID 3020 3072 BY SIMILARITY.  
FT DISULFID 3075 3075 BY SIMILARITY.  
FT CARBOHYD 151 151 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 237 237 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 564 564 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1170 1170 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1387 1387 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1622 1622 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1727 1727 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1847 1847 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1975 1975 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 1985 1985 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2093 2093 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2113 2113 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2161 2161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2276 2276 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2451 2451 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2647 2647 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2654 2654 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2663 2663 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2794 2794 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2810 2810 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2865 2865 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2929 2929 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 2964 2964 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 3028 3028 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 3028 3028 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT VARIANT 1288 R -> S.  
FT VARIANT 1305 T -> G.  
SQ SEQUENCE 3133 AA; 343350 MW; E5210D5D14A7B2B2 CRC64;

Query Match 16.0%; Score 79.5; DB 1; Length 3133;  
Best Local Similarity 28.4%; Pred. No. 2.6;  
Matches 27; Conservative 8; Mismatches 33; Indels 27; Gaps 6;  
QY 12 HAC-----IFQQLRCSSNPTLTCQRYCCEFFSLLHACP-----CLRSQPT 54  
DB 488 HACAHGVTFNMR-TNDLCPRMOCDEVCSNY-DSVCASCPVETDNLIIYAETARCEQDT 545  
QY 55 C-----OYCFHSEY-FDSLLHACPRATCQRYC 81  
DB 546 CVEGCKPKKSCPEGSYKNDSTECVPRAKCRVYC 580  
RESULT 14  
PKCS\_BRACL STANDARD; PRT; 1696 AA.  
AC 09NJ15; 09NJ16; 09NJ14;  
DT 16-OCT-2001 (Rel. 40; Created)  
DT 16-OCT-2001 (Rel. 40; Last sequence update)  
DE 16-OCT-2001 (Rel. 40; Last annotation update)  
DE Proprotein convertase subtilisin/kexin type 5 precursor (EC 3.4.21.-)  
DE (Proprotein convertase PC6-like) (apC6).  
GN PC6.  
OS Branchiostoma californiensis (California lancelet) (Amphioxus).  
OC Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;  
OC Branchiostoma.  
OX NCBI\_TaxID=7738;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS A; B AND C).  
RX MEDLINE=20175281; PubMed=10708868;  
RA OIiva A.A. Jr., Chan S.J., Steiner D.F.;  
RT "Evolution of the prothymone convertases: identification of a  
RT homologue of PC6 in the protochordate amphioxus.";  
RL Biochim. Biophys. Acta 1477:338-348(2000).  
CC -1- FUNCTION: LIKELY TO REPRESENT A WIDESPREAD ENDOPEPTASE ACTIVITY  
CC WITHIN THE CONSTITUTIVE AND REGULATED SECRETORY PATHWAY. CAPABLE  
CC OF CLEAVAGE AT THE RK(R)R CONSENSUS MOTIF (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: RELEASE OF MATURE PROTEINS FROM THEIR  
CC PROPEPTIDES BY CLEAVAGE OF ARG-XAA-YAA-ARG-I-ZAA BONDS, WHERE XAA  
CC CAN BE ANY AMINO ACID AND YAA IS ARG OR LYS.







GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 09:49:05 ; Search time 26.27 Seconds  
(Without alignments)  
533.407 Million cell updates/sec

Title: US-09-854-864-13  
Perfect score: 498  
Sequence: 1 CSQNEYFDSLHACIPCOLR.....SEYFDSLHACPPATCPQYC 81

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp.\_archaea:\*  
2: sp.\_bacteria:\*  
3: sp.\_fungi:\*  
4: sp.\_human:\*  
5: sp.\_invertebrate:\*  
6: sp.\_mammal:\*  
7: sp.\_mhc:\*  
8: sp.\_organelle:\*  
9: sp.\_phage:\*  
10: sp.\_plant:\*  
11: sp.\_protent:\*  
12: sp.\_virus:\*  
13: sp.\_vertebrate:\*  
14: sp.\_unclassified:\*  
15: sp.\_virus:\*  
16: sp.\_bacteriap:\*  
17: sp.\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	136	27.3	185	11	088472 mus musculu
2	100.5	20.2	1792	13	057484 galus gall
3	98	19.7	249	11	090823 mus musculu
4	98	19.7	249	11	090823 mus musculu
5	94.5	19.0	223	11	062220 mus musculu
6	94.5	19.0	223	11	064507 mus musculu
7	93.5	18.8	195	11	090141 mus musculu
8	93	18.7	293	4	014836 mus saplen
9	92.5	18.6	169	4	014564 mus saplen
10	92.5	18.6	227	11	090225 mus musculu
11	91.5	18.4	136	4	090825 mus saplen
12	91.5	18.4	186	4	090825 mus saplen
13	91.5	18.4	186	4	090825 mus saplen
14	91.5	18.4	191	11	090825 mus saplen
15	89	17.9	159	4	090825 mus saplen
16	89	17.9	1574	11	088281 ratius norv

17	88	17.7	166	4	090825 mus saplen
18	87.5	17.6	154	4	090825 mus saplen
19	87.5	17.6	174	4	090825 mus saplen
20	87.5	17.6	193	4	090825 mus saplen
21	87	17.5	159	4	090825 mus saplen
22	87	17.5	165	11	090825 mus saplen
23	87	17.5	195	11	090825 mus saplen
24	87	17.5	169	5	024301 mus musculu
25	86	17.3	188	11	070148 mus musculu
26	85.5	17.2	154	4	090825 mus saplen
27	85.5	17.2	177	11	090825 mus saplen
28	85.5	17.2	189	11	090825 mus saplen
29	85	17.1	202	11	091W93 mus musculu
30	85	17.1	325	10	094HS1 mus musculu
31	84.5	17.0	767	13	090825 mus saplen
32	84	16.9	195	4	090825 mus saplen
33	84	16.9	201	4	090825 mus saplen
34	84	16.9	210	4	090825 mus saplen
35	84	16.9	1671	5	090825 mus saplen
36	83.5	16.8	110	5	090825 mus saplen
37	83.5	16.8	130	11	090825 mus saplen
38	83.5	16.8	191	4	090825 mus saplen
39	83	16.7	175	4	090825 mus saplen
40	82	16.5	98	4	090825 mus saplen
41	82	16.5	126	6	028707 mus musculu
42	82	16.5	167	11	090825 mus saplen
43	82	16.5	168	11	090825 mus saplen
44	81.5	16.4	85	11	070555 mus musculu
45	81.5	16.4	429	13	091650 xenopus lae

## ALIGNMENTS

RESULT	ID	PRELIMINARY	PRT	185 AA.
088472	088472	088472		
AC	088472	088472		
DT	01-NOV-1998	(TREMUREL 08, Created)		
DT	01-NOV-1998	(TREMUREL 08, Last sequence update)		
DT	01-JUN-2001	(TREMUREL 17, Last annotation update)		
DE	B-CELL MATURATION PROTEIN (TUMOR NECROSIS FACTOR RECEPTOR SUPERFAMILY MEMBER 17).			
GN	TNFRSF17.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-BALB/C; TISSUE=SPLEN;			
RA	Madry C., Laabi Y., Callebaut I., Roussel J., Hatzoglu A.,			
RA	Leconiat M., Morton J.-P., Berger R., Tsapis A.,			
RT	"Murine BCMA: a new member of the Tumor Necrosis Factor Receptor Superfamily."			
RL	Submitted (APR-1998) to the EMBL/Genbank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-C57BL/6J; TISSUE=COLON.			
RX	MEDLINE=21085660; PubMed=11217851;			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA	Aizawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,			
RA	Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito R.,			
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,			
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,			
RA	Schirral L.M., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,			
RA	Sakai K., Okido T., Futuno M., Aono H., Baldarelli R., Barz G.,			
RA	Blake J., Boffelli D., Bojunga N., Carinica P., de Bonaldo M.F.,			
RA	Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,			
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,			
RA	Lyons P., Marchionni L., Mashima J., Mazzaletti J., Mombaerts P.,			

RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AF061505; AAC23799.1; -  
 DR EMBL: AK020247; BAB32038.1; -  
 DR MGD: MGI:1343050; Tinf17.  
 SO SEQUENCE 185 AA; 20442 MW; 8806352BAFD26A8E CRC64;

Query Match 27.3%; Score 136; DB 11; Length 185;  
 Best Local Similarity 70.6%; Pred. No. 1.7e-09;  
 Matches 24; Conservative 2; Mismatches 6; Indels 2; Gaps 1;

OY 1 CSONEYFDSLHACIPQOLRCSSNTPPLTCORC 34  
 DB 5 CPHESEFDSLHACRPHLRCSN-PPATCOPYC 36

RESULT 2

057484 PRELIMINARY; PRT; 1792 AA.

AC 057484;

DT 01-JUN-1998 (TREMBLrel. 06, Created)

DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE LAMININ BETA 2-LIKE CHAIN.

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI\_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-93015947; PubMed-1400373;

RA O'Rear J.J.;

RT "A novel laminin B1 chain variant in avian eye.";

RL J. Biol. Chem. 267:2055-2055(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE-96209634; PubMed-9550264;

RA Liu J., Swadlow S., Xie W., Brewton R.G., Mayne R.;

RT "Primary structure and expression of a chicken laminin beta chain:

RT evidence for four beta chains in birds.";

RL Matrix Biol. 16:471-481(1998).

DR EMBL: AF038555; AAB92586.1; -

DR HSPF: P02468; IKLO.

DR InterPro: IPR002106; AA\_trna\_ligase\_II.

DR InterPro: IPR000561; EGF-like.

DR InterPro: IPR002049; Laminin\_EGF.

DR InterPro: IPR001886; LamNT.

DR Pfam: PF00053; Laminin\_EGF\_13.

DR Pfam: PF00055; Laminin\_Nterm; 1.

DR PRINTS: PR00011; EGF\_LAMININ.

DR PRODOM: PD002082; LamNT; 1.

DR SMART: SM00180; EGF\_Lam; 13.

DR SMART: SM00136; LamNT; 1.

DR PROSITE: PS00039; AA-trna\_ligase\_II\_2; UNKNOWN\_1.

DR PROSITE: PS00022; EGF\_1; UNKNOWN\_10.

DR PROSITE: PS01186; EGF\_2; 2.

DR PROSITE: PS01248; LAMININ\_TYPE\_EGF; 12.

DR EGF-like domain; Glycoprotein; Laminin EGF-like domain; Repeat.

SO SEQUENCE 1792 AA; 195723 MW; 4A4CBED0206F6A600 CRC64;

DB 1017 CDPGYGDAMRHSC-----RRSCNTLGTBPNTGPOCCOCDORSOGCHLPHEGSCDR 1072  
 OY 50 CSP-----PTCYCCFHESEFDSLHACPPATCOPYC 81  
 DB 1073 CSPENWNLGSGGCEPCACHPQH--SLSPACNOFTQCSC 1110

RESULT 3

ID 09DB23 PRELIMINARY; PRT; 249 AA.

AC 09DB23;

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE 1200009E08RIK PROTEIN.

GN TNFRSF13B OR 1200009E08RIK.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX STRAIN=C57BL/6J; TISSUE=LUNG;

RX MEDLINE-21085660; PubMed-11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa K., Nishi K., Kiyosawa H., Kondo S., Yamana K.,

RA Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schiml L.M., Staubl F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Whitaker C., Wilming L.,

RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RA Hayashizaki Y., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL: AK004668; BAB23457.1; -

DR MGD: MGI:1889411; Tnf13b.

SO SEQUENCE 249 AA; 26981 MW; 6F4290D719FEA037 CRC64;

Query Match 19.7%; Score 98; DB 11; Length 249;  
 Best Local Similarity 25.6%; Pred. No. 0.00012;  
 Matches 23; Conservative 12; Mismatches 27; Indels 28; Gaps 4;

OY 1 CSONEYFDSLHACIPQOLRCSSNTPPLTCORCCEYFDSLHACPLGSPPTCYCCF 60  
 DB 6 CPKDYWDSSRSCVSCALTCORS-----QRTCTDE-----KFTNCKRE----- 46

OY 61 HSEYFDSLHAC-----PPATCOPYC 81  
 DB 47 QGRYDHLGACVSCDSTCTGHPQCAHRC 76

RESULT 4

ID 09ET35 PRELIMINARY; PRT; 249 AA.

AC 09ET35;

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE TACI PROTEIN.

GN TNFRSF13B.

OS Mus musculus (Mouse).

OY 1 CSONEYFDSLHACIPQOLRCSSN--TPPLTC--QRYCCFDSLHACP-----CLR 49

```

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. PubMed=10881172;
RX MEDLINE=2117254;
RA Yan M., Marsters S.A., Grewal I.S., Wang H., Ashkenazi A., Dixit V.M.;
RT "Identification of a receptor for Blys demonstrates a crucial role in
RT humoral immunity";
RL Nat. Immunol. 1:37-41(2000).
DR EMBL: AF257673; AKG00081.1; -.
DR MGI: M37759; AAA40106.1; -.
SQ SEQUENCE 249 AA; 26947 MW; CB2F2D61C2931D81 CRC64;

Query Match
Best Local Similarity 19.7%; Score 98; DB 11; Length 249;
Matches 23; Conservative 12; Mismatches 27; Indels 28; Gaps 4;

OY 1 CSQNEFYDSLHACIPCOLRCSSTNPPLTCQRYCCEYFDSLHACPLRCSPPTCQYCCF 60
DB 6 CRRDQYWDSSRKSCVSCALTCSSRS-----QRTCTDF-----CKFNGCRKE----- 46
OY 61 HSEYFDSLHAC-----PPATCQPYC 81
DB 47 OGRYVDHLGACYSQDSTQTHPOCAHFC 76

RESULT 5
O62220 PRELIMINARY; PRT; 223 AA.
AC O62220;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SERINE 2 ULTRA HIGH SULFUR PROTEIN.
GN KRTAP5-4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=91065960; PubMed=2250030;
RA Wood L., Mills M., Hatzenbuehler N., Vogeli G.;
RT "Additions and Corrections: Serine-rich ultra high sulfur protein gene
RT expression in murine hair and skin during the hair cycle.";
RL J. Biol. Chem. 265:4024-4024(1991).
DR EMBL: M37760; AAA40107.1; -.
DR HSSP: P01064; IP12.
DR MGI: M37759; AAA40106.1; -.
DR MGD: MGI:1354732; Krtap5-4.
DR InterPro: IPR001450; 4Fe4S_Ferredoxin.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR001305; DnaJ_CXXCXXG.
DR InterPro: IPR001007; VMFC.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
DR PROSITE: PS00177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE: PS00637; DNAJ_CXXCXXG; UNKNOWN_1.
DR PROSITE: PS01208; VMFC; UNKNOWN_2.
DR PROSITE: PS01208; VMFC; UNKNOWN_2.
SQ SEQUENCE 223 AA; 21442 MW; C654BDB9FD08C59A CRC64;

Query Match
Best Local Similarity 19.0%; Score 94.5; DB 11; Length 223;
Matches 24; Conservative 3; Mismatches 27; Indels 27; Gaps 5;

OY 1 CSQNEFYDSLHACIPCOLRCSSTNPPLTCQRYCCEYFDSLHACPLRCSPPTCQYCCF 60
DB 61 HSEYFDSLHACPPATCQPYC 81

```

```

DB 150 CCOS-----CCKPC---CSSGCGSSCCOSSCC-----PCC-COSSCCPCC 189
OY 61 HSEYFDSLHACPPATCQPYC 81
DB 190 QSS-----CCKPCCOSSC 203

RESULT 6
O64507 PRELIMINARY; PRT; 230 AA.
AC O64507;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SERINE 1 ULTRA HIGH SULFUR PROTEIN.
GN KRTAP5-1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BALB/C;
RX MEDLINE=91065960; PubMed=2250030;
RA Wood L., Mills M., Hatzenbuehler N., Vogeli G.;
RT "Serine-rich ultra high sulfur protein gene expression in murine hair
RT and skin during the hair cycle.";
RL J. Biol. Chem. 265:21375-21380(1990).
DR EMBL: M37759; AAA40106.1; -.
DR HSSP: P01064; IP12.
DR MGI: MGI:1354732; Krtap5-1.
DR InterPro: IPR001450; 4Fe4S_Ferredoxin.
DR InterPro: IPR000020; Anaphylatoxin.
DR InterPro: IPR001305; DnaJ_CXXCXXG.
DR InterPro: IPR001007; VMFC.
DR PROSITE: PS00198; 4FE4S_FERREDOXIN; UNKNOWN_1.
DR PROSITE: PS00177; ANAPHYLATOXIN_1; UNKNOWN_1.
DR PROSITE: PS00637; DNAJ_CXXCXXG; UNKNOWN_1.
DR PROSITE: PS01208; VMFC; UNKNOWN_3.
DR PROSITE: PS01208; VMFC; UNKNOWN_3.
SQ SEQUENCE 230 AA; 21781 MW; 6CC50B41B2137C23 CRC64;

Query Match
Best Local Similarity 19.0%; Score 94.5; DB 11; Length 230;
Matches 24; Conservative 3; Mismatches 27; Indels 27; Gaps 5;

OY 1 CSQNEFYDSLHACIPCOLRCSSTNPPLTCQRYCCEYFDSLHACPLRCSPPTCQYCCF 60
DB 149 CCQSS-----CCKPC---CSSGCGSSCCOSSCC-----PCC-COSSCCPCC 188
OY 61 HSEYFDSLHACPPATCQPYC 81
DB 189 QSS-----CCKPCCOSSC 202

RESULT 7
O9D141 PRELIMINARY; PRT; 195 AA.
AC O9D141;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 18 DAYS EMBRYO CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY.
DE CLONE:1110030N11, FULL INSERT SEQUENCE.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
RX MEDLINE=21085650; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

```

Query Match	18.8%	Score 93.5	DB 11	length 195
Best Local Similarity	33.3%	Pred. No. 0.00035		
Matches	26	Conservative	3	Mismatches 24; Indels 25; Gaps 6
OY	14	CIP---COLARC-----SSNPPLTCORCYCCCFPSLHAC--PCLRCSPPTGOCYCCFHSE	63	
Db	89	CQPSCCGSSCCQPCSCGSSCCQPCSCGSSCCQPCRCISSCCPC--CRPSCGSSC-----	142	
OY	64	YFDSLHACPPATCOPYC	81	
Db	143	-----CKPC-CQPC	151	

```

Query Match      18.7%  Score 93  DB 4:  Length 293;
Best Local Similarity 28.4%  Pred. No. 0.0006
Matches 21:  Conservative 13:  Mismatches 22:  Indels 18:  Gaps 5:

      K_CSN0NEYFDLSLHACIPCOLRCSSNPTPLTCORYC---C-----EYFDSLHACPOLRCS 51
      +-+:+:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:+++:
34  QPEQYDMPPLGLGCMSCCTTCICNHOS-QRTGCAFCFSLSCRKQEGFVPHLLR--QCISCA 90

```

Query Match	18.6%	Score 92.5	DB 4	Length 169
Best Local Similarity	30.0%	Pred. No. 0.00041		
Matches	27	Conservative	5	Mismatches 37; Indels 21; Gaps 5;

  

QY	1	CSOMEYFDLSLHACIPQLRCSSMTPLTCORRCCEYFDLSLHAC-PLLR-----CS	51
Db	77	CSQ-----CSCCKPFC--CCSGCGSCCCSCCKPFCSCGSCCKPFCSSSGKSSCCQ	127
QY	52	PTPCQYCCFHSSEYFDLSLHACPPATCPQYC	81
Db	128	SSCKPCCSSSGCGSS--CCGSSCKPFC	154

RESULT	10		
09D225			
ID	09D225		
AC	09D225	PRELIMINARY;	PRN; 227 AA.
DT	01-JUN-2001 (Tremblrel. 17, Created)		
DT	01-JUN-2001 (Tremblrel. 17, Last sequence update)		
DT	01-DEC-2001 (Tremblrel. 19, Last annotation update)		
DE	A030009A09RIK PROTEIN.		
GN	A030009A09RIK.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=C57BL/6J; TISSUE=SKIN;		
RX	MEDLINE=1085660; Pubmed=11217851;		
RA	Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,		
RA	Arakawa T., Hara A., Fukunishi Y., Komio H., Adachi J., Fukuda S.,		
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,		
RA	Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,		
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,		
RA	Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,		

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schimpl L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsch G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.T., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyszewski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,  
 RA Hayashizaki Y.,  
 RT "Functional annotation of a full-length mouse cDNA collection."  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK020700; BAB32181.1; -  
 DR HSSP: Q46655; 1CJH.  
 DR MGD: MGI:1925013; A030009A09Rik.  
 DR InterPro: IPR002494; Keratin\_B2.  
 DR Pfam: PF01500; Keratin\_B2.  
 DR PROSITE: PS01208; VMFC; UNKNOWN\_1.  
 SQ SEQUENCE 227 AA; 23407 MW; 9A89B93A01E13E45 CRC64;

Query Match 18.6%; Score 92.5; DB 11; Length 227;  
 Best Local Similarity 25.8%; Pred. No. 0.00054;  
 Matches 23; Conservative 4; Mismatches 33; Indels 29; Gaps 4;  
 QY 14 CIP--COLGSSNTPLT-----CQRCCEYFDSLHACPLCRSP 53  
 DB 58 CVPSCCPSCCAPCLTLCITPVSVCSSCCSCCTPSCCQSSCCPACCTCSCPCPS 117  
 QY 54 TC-QYCFHSEYFDSLHACPPATCPYC 81  
 DB 118 CCVPVCTPV-----CCTPVCKKPYC 138

RESULT 11  
 Q9BYR5 PRELIMINARY; PRT; 136 AA.  
 AC Q9BYR5:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE KERATIN ASSOCIATED PROTEIN 4.2.  
 GN KRTAP4.2.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SCALP;  
 RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.,  
 RT "Characterization of a cluster of human high/ultrahigh keratin  
 RT associated proteins on chromosome 17q12-21."  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ406934; CAC27573.1; -  
 DR InterPro: IPR002494; Keratin\_B2.  
 DR InterPro: IPR001368; TNER\_C6.  
 DR Pfam: PF01500; Keratin\_B2; 1.  
 DR PROSITE: PS00652; TNER\_NGFR\_1; UNKNOWN\_1.  
 SQ SEQUENCE 136 AA; 14402 MW; 51E514863925E02E CRC64;

Query Match 18.4%; Score 91.5; DB 4; Length 136;  
 Best Local Similarity 29.7%; Pred. No. 0.00045;  
 Matches 27; Conservative 5; Mismatches 38; Indels 21; Gaps 5;  
 QY 10 LTHACIP--COLGSSNTPLT-----CQRCCEYFDSLHAC--PCLR---C 50  
 DB 17 LENCPCPSCCTTCCTTCRCRPSVCSSCCRPCCQSSCCQCPICSCPCCTTCCTCC 76

QY 51 SPTCYCFHSEYFDSLHACPPATCPYC 81  
 DB 77 RPSVCSSCRPPCCQSV--CCOPTCCRPSC 105

RESULT 12  
 Q9BYR2 PRELIMINARY; PRT; 186 AA.  
 AC Q9BYR2:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE KERATIN ASSOCIATED PROTEIN 4.5.  
 GN KRTAP4.5.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SCALP;  
 RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.,  
 RT "Characterization of a cluster of human high/ultrahigh keratin  
 RT associated proteins on chromosome 17q12-21."  
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AJ406937; CAC27576.1; -  
 DR InterPro: IPR002494; Keratin\_B2.  
 DR InterPro: IPR001304; Iectin\_C.  
 DR InterPro: IPR001368; TNER\_C6.  
 DR InterPro: IPR001007; VMFC.  
 DR Pfam: PF01500; Keratin\_B2; 1.  
 DR PROSITE: PS00615; C-TYPE\_LECTIN\_1; UNKNOWN\_1.  
 DR PROSITE: PS00652; TNER\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE: PS01208; VMFC; UNKNOWN\_2.  
 SQ SEQUENCE 186 AA; 19916 MW; 034D9C734304F63A CRC64;

Query Match 18.4%; Score 91.5; DB 4; Length 186;  
 Best Local Similarity 30.5%; Pred. No. 0.0006;  
 Matches 25; Conservative 6; Mismatches 28; Indels 23; Gaps 6;  
 QY 10 LTHACIP--COLGSSNT--PPLTCQRCCEYFDSLHACPLCRSPPTC-----OYCC 59  
 DB 17 LENCPCPSCCTTCCTTCRCRPSVCSSCCRPCCQ--SVCYQPTC--CHPSCCISCCRPVCC 71  
 QY 60 FHSYFDSLHACPPATCPYC 81  
 DB 72 ESS-----CRRPCCCRPSC 85

RESULT 13  
 Q64526 PRELIMINARY; PRT; 136 AA.  
 AC Q64526:  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ULTRA-HIGH SULPHUR KERATIN.  
 GN KRTAP9-1.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=89140394; PubMed=2465353;  
 RA McNab A.R., Wood L., Theriault N., Gierman T., Vogel G.,  
 RT "An ultra-high sulfur keratin gene is expressed specifically during  
 RT hair growth."  
 RL J. Invest. Dermatol. 92:263-266(1989).  
 CC -1- SIMILARITY: TO THE PLANT THIONIN FAMILY.  
 DR EMBL: M27685; AAA81560.1; -  
 DR MGD: MGI:1309997; Krtap9-1.

DR InterPro: IPR002494; Keratin\_B2.  
 DR InterPro: IPR001010; Thionin.  
 DR InterPro: IPR001368; TNFR\_C6.  
 DR InterPro: IPR001007; WMF.  
 DR Pfam: PF01500; Keratin\_B2; 1.  
 DR PROSITE: PS00271; THIONIN; 1.  
 DR PROSITE: PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 DR PROSITE: PS01208; WMF; UNKNOWN\_1.  
 SO SEQUENCE 186 AA; 19658 MW; 61D6BFDAB72CFEB4 CRC64;

Query Match 18.4%; Score 91.5; DB 11; Length 186;  
 Best Local Similarity 34.7%; Pred. No. 0.0006;  
 Matches 25; Conservative 4; Mismatches 24; Indels 19; Gaps 6;

OY 14 CIP-COLRC--SSNPPITCQRYCCFEDSLHAC--PCLRCSPPTCQYCFHSEYFDSL 69  
 DB 85 CQPCCPSCCQSCCQPCPCSCCQPCRCISSCCQPC--CRPSCCQSSC----- 132  
 OY 70 HACPPATCQPYC 81  
 DB 133 --CRPC-CQPC 141

RESULT 14  
 OY9D3H7 PRELIMINARY; PRT; 191 AA.  
 AC OY9D3H7;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE 10 DAYS NEONATE HEAD CDNA, RIKEN FULL-LENGTH ENRICHED LIBRARY,  
 DE CLONE:5530401102, FULL INSERT SEQUENCE.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=HEAD;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aikawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
 RA Aikawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,  
 RA Saito T., Okazaki Y., Gojodori T., Bono H., Kasukawa T., Saito K.,  
 RA Kadota K., Matsuda H.A., Ashburner T., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schirini L.M., Straubli F., Suzuki R., Tomita M., Wagner L., Mashio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,  
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,  
 RA Hayashizaki Y.;  
 RT "Functional annotation of a full-length mouse cDNA collection";  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK017437; BAB30743.1.  
 DR InterPro: IPR002494; Keratin\_B2.  
 DR InterPro: IPR001010; Thionin.  
 DR Pfam: PF01500; Keratin\_B2; 1.  
 DR PROSITE: PS00271; THIONIN; UNKNOWN\_1.  
 SO SEQUENCE 191 AA; 20088 MW; 522B841DC9A8A9D5 CRC64;

Query Match 18.4%; Score 91.5; DB 11; Length 191;  
 Best Local Similarity 32.9%; Pred. No. 0.00062;  
 Matches 23; Conservative 3; Mismatches 21; Indels 23; Gaps 5;

OY 16 PQLRCSSNTP---PLTCQRYCCFEDSLHACPCRLCSPPTCQ-YCFHSEYFDSLHA 71  
 DB 63 PCVSSCCRTPCQPCQCCVSSCCQ-----PC--CQPCSCQSSCCQPS-----C 103  
 OY 72 CPPATCQPYC 81  
 DB 104 CQPCSCQPC 113

RESULT 15  
 OY9BYO3 PRELIMINARY; PRT; 159 AA.  
 AC OY9BYO3;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE KERATIN ASSOCIATED PROTEIN 9.3.  
 GN KRTP9.3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=SCALP;  
 RA Rogers M.A., Langbein L., Winter H., Ehmann C., Korn B., Schweizer J.,  
 RT "Characterization of a cluster of human high/ ultrahigh keratin  
 RT associated proteins on chromosome 17q12-21.";  
 RL Submitted (Oct-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AJ406947; CAC27586.1.  
 DR InterPro: IPR002494; Keratin\_B2.  
 DR InterPro: IPR001368; TNFR\_C6.  
 DR Pfam: PF01500; Keratin\_B2; 1.  
 DR PROSITE: PS00652; TNFR\_NGFR\_1; UNKNOWN\_1.  
 SO SEQUENCE 159 AA; 16853 MW; 375C1E52EECD68 CRC64;

Query Match 17.9%; Score 89; DB 4; Length 159;  
 Best Local Similarity 29.5%; Pred. No. 0.0011;  
 Matches 26; Conservative 4; Mismatches 20; Indels 38; Gaps 7;

OY 10 LHAACP-COLRC-----SSNPPITCQRYCCFEDSLHAC--PCLRCSP 53  
 DB 1 MHCSPCCQPTCCRTCTQPTVTTCSTP--CQPCSC-----VSSCCQPC--CHFT 50  
 OY 54 TCQYCCFHSYFDSLHACPPATCQPYC 81  
 DB 51 CCQNTC-----CRTCCQPC 66

Search completed: May 15, 2002, 09:53:06  
 Job time: 241 sec







FT		/note= "Derived from human BCMA"
XX		
FN		WO200124811-A1.
XX		
PD		12-APR-2001.
XX		
PF		05-OCT-2000; 2000WO-US27579.
XX		
PR		06-OCT-1999; 99US-0157933.
XX		11-FEB-2000; 2000US-0181807.
PR		30-JUN-2000; 2000US-0215688.
XX		
PA	(BIOJ ) BIOGEN INC.	
XX	(ABOT-) APOTECH R & D SA.	
PZ	Schneider P, Thompson J, Cachero T, Ambrose C, Rennett P;	
DR	WPI: 2001-266242/27.	
XX	N-PSDB: AAD03847.	
PT	Treating a mammal for a condition associated with undesired cell	
PT	proliferation such as cancer or carcinoma, comprises administering a	
PT	composition comprising A Proliferation Inducing Ligand Receptor	
PT	(APRIL-R) antagonist -	
XX		
PS	Example 1; Fig 3b; 85pp; English.	
XX		
CC	The invention relates to a method of treating a mammal for a condition	
CC	associated with undesired cell proliferation such as cancer or	
CC	carcinoma. The method involves administering a composition comprising	
CC	A Proliferation Inducing Ligand Receptor (APRIL-R) also referred as	
CC	B cell maturation Protein (BCM or BCMA) antagonist that antagonises the	
CC	interaction between APRIL and its cognate receptor(s). This method is	
CC	useful for treating undesired cell proliferation such as cancer or	
CC	carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,	
CC	prostate carcinoma, and other carcinomas whose proliferation is modulated	
CC	by APRIL. It is also useful for treating autoimmune diseases (Grave's	
CC	disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular	
CC	diseases, renal disorders, B-cell lympho-proliferative disorders,	
CC	Immunosuppressive diseases, organ transplantation, inflammation and	
CC	human immunodeficiency virus (HIV), and for treating, suppressing or	
CC	altering an immune response involving a signalling pathway between	
CC	APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.	
CC	The present sequence is a fusion construct containing human APRIL-R	
CC	also referred as BCMA or BCM protein, FC region of human Immunoglobulin	
CC	G (IgG) and a signal sequence from murine Ig kappa cDNA.	
XX		
SO	Sequence 302 AA;	
	Query Match 41.2%; Score 205; DB 22; Length 302;	
	Best Local Similarity 67.8%; Pred. No. 1.9e-10;	
	Matches 40; Conservative 2; Mismatches 9; Indels 8; Gaps 3.	
OY	1 CSQNEFDSLHACIPCOLRCSSNPPTELCQRYC-C EYFDSL-----LHACPCLRCSP 53	
Dd	31 csqneyfslshacipcolrcssnpplteclcqrycnsavtvsykydkhtcp--pcpap 87	
RESULT 2	.	
AAAB60699		
ID	AAAB60699 standard; Protein; 302 AA.	
XX		
AC	AAAB60699;	
XX		
DT	22-MAY-2001 (first entry)	
DE	Mouse Igg signal/human BAFF-R/human Igg Fc fusion protein, BAFF-R-Fc.	
XX		
KM	Human BAFF-R: BAFF receptor; TNF family; immunoregulatory agent;	
KW	immune-related disorder; B-cell growth inhibitor;	
KW	B-cell maturation inhibitor; immunoglobulin production inhibitor;	
KW	autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;	

KX	renal disorder; immunosuppressive disorder; HIV infection;
KW	organ transplantation; antinflammatory; systemic lupus erythematosus;
KM	autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;
KV	B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;
KX	lymphoma; gene therapy; cancer; tumour; IgG Fc; fusion construct.
OS	Chimeric - Homo sapiens.
XX	Chimeric - Mus sp.
PN	WO200112812-A2.
PD	
XX	22-FEB-2001.
PF	16-AUG-2000; 2000MO-US22507.
XX	
PR	17-AUG-1999; 99US-0149378.
PR	11-FEB-2000; 2000US-0181684.
PR	18-FEB-2000; 2000US-0183536.
XX	
PA	(BIOJ ) BIOGEN INC.
PP	(APOT ) APOTEC R & D SA.
Pt	
Pt	Mackay F, Browning J, Ambrose C, Tschopp J, Schneider P,
Pt	Thompson J;
DR	
XX	WPI, 2001-202866/20.
XX	N-PSDB; AAF59999.
Pt	
Pt	Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell
Pt	lymphoproliferative disorder by administering BAFF-receptor
Pt	polypeptide, chimeric molecule comprising receptor or anti-BAFF-R
PT	antibody homolog -
PS	
XX	Example 4; Fig 2; 59pp; English.
CC	The invention relates to the use of a BAFF receptor (BAFF-R, also known
CC	as BCMa) protein, or a BAFF-R fusion protein as an agent for the
CC	treatment of a variety of immune-related disorders. BAFF-R is a member of
CC	the TNF (tumour necrosis factor) family, acting as an immunoregulatory
CC	agent, and also plays a role in the development of hypertension and
CC	related disorders. BAFF-R, fusion proteins containing it, and BAFF-R-
CC	specific antibodies can be used for inhibiting B-cell growth, dendritic
CC	cell-induced B-cell growth and maturation, and immunoglobulin production,
CC	and in the treatment of autoimmune disorders, B-cell lymphoproliferative
CC	disorders, hypertension and renal disorders. The BAFF-R proteins may also
CC	be used in the treatment of immunosuppressive disorders and HIV
CC	infection, and in patients undergoing organ transplantation. The BAFF-R
CC	proteins or BAFF-R specific antibodies may be used for treating,
CC	suppressing or altering an immune response involving a signalling pathway
CC	between BAFF-R and BAFF, thereby inhibiting inflammation. Since BAFF-R
CC	inhibits B-cell growth and maturation it is useful for treating diseases
CC	such as systemic lupus erythematosus, autoimmune haemolytic anaemia,
CC	grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly
CC	progressive glomerulonephritis, and lymphomas. Nucleic acids encoding
CC	human BAFF-R may be used in gene therapy to treat tumours, lymphomas,
CC	autoimmune disorders and inherited B-cell-associated disorders. The
CC	present sequence represents the BAFF-R fusion protein BAFF-R-Fc,
CC	comprising a mouse IgG-kappa signal sequence, residues 1-153
CC	of human BAFF-R and a human IgG Fc sequence.
XX	
SQ	Sequence 302 AA:
Query Match	41.2%; Score 205; DB 22; Length 302;
Best Local Similarity	67.8%; Pred. No. 1,9e+10;
Matches 40; Conservative 2; Mismatches 9;	Indels 8; Gaps 3
1 CSQNFYPSLHACIPQCDCSSNTPPLTLCQRYC-C-EYFPDSL-----LHACPCLRCSP 53	:
db 31 csqneyfdfsllhacipqcqrccsnptlctcqrycnasvtnskvydkvthtcp--pcpap 87	:

AA08843	
ID	AA08843 standard; peptide; 184 AA.
XX	
AC	AA08843;
XX	
DT	02-JAN-2001 (first entry)
XX	
DE	Amino acid sequence of human.
XX	
XX	BCMA: necrosis factor-kB activator; NF-kB; gene expression; cancer;
KW	anti-cell death gene; apoptosis; viral infection; inflammatory response;
KW	rheumatoid arthritis; inflammatory bowel disease; septic shock.
XX	
OS	Homo sapiens.
XX	
FT	Key
FT	Domain
FT	Location/Qualifiers
FT	57..77
FT	/note="putative transmembrane domain"
XX	
PN	WO200050633-A1.
XX	
PD	31-AUG-2000.
XX	
PF	24-FEB-2000; 2000WO-US04925.
XX	
PR	24-FEB-1999; 99US-0121485.
XX	
PA	(GCHO) GEN HOSPITAL CORP.
XX	
PI	Seed B, Ting A;
XX	
DR	WPI; 2000-558405/B1.
XX	
PT	Identifying a modulator of gene expression for drug designing, by
PT	contacting a compound library with a cell expressing an anti-cell death
PT	gene and reporter gene, and determining alteration in reporter gene
PT	expression
XX	
PS	Claim 32; Fig 7A; 53pp; English.
XX	
XX	The present sequence represents a BCMA (not defined) polypeptide. BCMA
CC	is a necrosis factor (NF)-kB activator. The method of the invention is
CC	used to identify compounds which modulate BCMA activity (and thus NF-kB
CC	activity). The specification describes a method of identifying a
CC	polypeptide which increases gene expression from a promoter. The method
CC	involves contacting a library of with a cell which expresses a
CC	recombinant anti-cell death gene and a reporter gene operably linked to
CC	the promoter, and then determining whether the expression of the
CC	reporter gene is altered as a result of contact with library. The method
CC	is useful for identifying polypeptides which increase or decrease gene
CC	expression from a promoter. The BCMA polypeptide or nucleic acid are
CC	useful for preparing a pharmaceutical composition for treating cancer,
CC	apoptosis, viral infections, inflammatory response, such as rheumatoid
CC	arthritis, inflammatory bowel disease or septic shock. BCMA is useful for
CC	identifying compounds that modulate NF-kB expression and thus for drug
CC	designing.
XX	
XX	Sequence 184 AA:
SO	
QY	Query Match
	Best Local Similarity 40.4%; Score 201; DB 21; Length 184;
	Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 CSONEFDSLHACIPCOLRCSNTPPLTCQRC 34
	8 csqneyfdsllhacipqrlrcssnptpctqrc 41
RESULT	4
AA094001	
ID	AA094001 standard; Protein; 184 AA.
XX	

AC	AA094001;
XX	
XX	20-OCT-2000 (first entry)
DT	
XX	
DE	A human BCMA protein, a B cell protein related to TACI.
XX	
KW	Human; BR43x2; TACI receptor; extracellular domain; BCMA; B cell protein;
KW	transmembrane activator and CAML-interactor; tumour necrosis factor; TNF;
KW	ztnf4 activity; antibody production; autoimmune disease; amyloidosis;
KW	systemic lupus erythematosus; myasthenia gravis; multiple sclerosis;
KW	rheumatoid arthritis; asthma; bronchitis; emphysema; pyelonephritis;
KW	end stage renal failure; glomerulonephritis; vasculitis; nephritis;
KW	renal neoplasms; multiple myeloma; lymphoma; light chain neuropathy;
KW	immune response; immunosuppression; graft rejection; joint pain;
KW	graft versus host disease; inflammation; swelling; anaemia; septic shock;
KW	insulin dependent diabetes mellitus; Crohn's disease; hypertension;
KW	renal artery stenosis; occlusion; cholesterol; renal emboli.
XX	
OS	Homo sapiens.
XX	
PN	WO200040716-A2.
XX	
PD	13-JUL-2000.
XX	
PF	07-JAN-2000; 2000WO-US00396.
XX	
PR	07-JAN-1999; 99US-0226533.
XX	
PA	(ZYMO) ZYMOGENETICS INC.
XX	
PI	Gross JA, Xu W, Madden K, Yee DP;
XX	
DR	WPI; 2000-452538/39.
XX	
PT	Inhibiting ztnf4 activity in a mammal, to treat autoimmune diseases,
PT	renal disease, graft versus host disease, and inflammation, comprises
PT	administering a BR43x2, TACI or BCMA extracellular domain polypeptide
XX	
PS	Disclosure; Page 152; 175pp; English.
XX	
XX	The present sequence represents a human BCMA protein, a B cell protein
CC	related to transmembrane activator and CAML-interactor (TACI) receptor.
CC	TACI is a tumour necrosis factor (TNF) receptor. The extracellular
CC	domains of BR43x2 (an isoform of TACI), TACI or BCMA (a related B cell
CC	protein) receptor contain a cysteine rich domain, and are used for
CC	inhibiting ztnf4 activity. ztnf4 is a TNF ligand. They may also be used
CC	for inhibiting BR43x2, TACI or BCMA receptor-ligand engagement associated
CC	with activated or resting B lymphocytes, effector T-cells, or with
CC	autoimmune disease selected from systemic lupus erythematosus, myasthenia
CC	gravis, multiple sclerosis and rheumatoid arthritis. The ztnf4 activity
CC	and BR43x2, TACI or BCMA receptor-ligand engagement is associated with
CC	asthma, bronchitis, emphysema, end stage renal failure,
CC	glomerulonephritis, vasculitis, nephritis, pyelonephritis, renal
CC	neoplasms, multiple myelomas, lymphomas, light chain neuropathy,
CC	amyloidosis, moderating immune response, immunosuppression, graft
CC	rejection, graft versus host disease, inflammation, insulin dependent
CC	diabetes mellitus, Crohn's disease, joint pain, swelling, anaemia, or
CC	septic shock. BR43x2, TACI, and BCMA polypeptides, fusions, antibodies,
CC	agonists or antagonists can be used to treat hypertension, renal artery
CC	stenosis, or occlusion, and cholesterol or renal emboli.
XX	
XX	Sequence 184 AA:
SO	
QY	Query Match
	Best Local Similarity 40.4%; Score 201; DB 21; Length 184;
	Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db	1 CSONEFDSLHACIPCOLRCSNTPPLTCQRC 34
	8 csqneyfdsllhacipqrlrcssnptpctqrc 41

```

RESULT      5
AAE09241    AAE09241 standard; Protein; 184 AA.
XX          AAE09241;
AC           Human TNF, tumour necrosis factor; TALL-1; APRIL; TNF receptor;
XX          TNFR; TACI; BCMA; therapy; cancer; leukaemia; myeloma; lymphoma;
KW          autoimmune disease; rheumatoid arthritis; multiple sclerosis;
KW          psoriasis.
XX
OS          Homo sapiens.
XX
PN          WO200160397-A1.
XX
PD          23-AUG-2001.
XX
PE          28-NOV-2000; 2000WO-US32378.
XX
PR          16-FEB-2000; 2000US-0182938.
XX          22-AUG-2000; 2000US-0226986.
PA          (GETH ) GENENTECH INC..
XX
PI          Ashkenazi AJ, Dodge KH, Grewal I, Kim KJ, Marsters SA, Pitti RM;
PI          Yan M;
DR          WPI: 2001-541628/60.;
DR          N-PSDB: AAD15902.
XX
PT          Inhibiting or neutralizing TALL-1 or APRIL polypeptide biological
PT          activity, for treating autoimmune disorders and cancer, comprises
PT          exposing the cells to TALL-1 or APRIL polypeptide agonists or
PT          antagonists -
XX
PS          Example 2; Fig 2; 160pp; English.
XX
CC          The invention relates to methods of using one or more agonists or
CC          antagonists to modulate the activity of the members of TNF (tumour
CC          necrosis factor) especially TALL-1, APRIL and TNF receptor (TNFR)
CC          e.g. TACI or BCMA. The method is useful for treating pathological
CC          conditions or diseases associated with increased TALL-1 and APRIL
CC          expression or activity. TALL-1 and APRIL antagonists are used to
CC          block the interaction between APRIL and TALL-1 with TACI or BCMA.
CC          They are useful for treating a mammal suffering from cancer such
CC          as leukaemia, lymphoma, myeloma, cancers of lung and colon and
CC          autoimmune diseases e.g. rheumatoid arthritis, multiple sclerosis,
CC          psoriasis and lupus erythematosus. The present sequence is human
CC          BCMA protein.
XX
SQ          Sequence      184 AA;

Query Match              40.4%; Score 201; DB 22; Length 184;
Best Local Similarity   100.0%; Pred. No. 2.8e-10;
Matches     34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 CSQNEYFDLSLHACTPCQLRCSSNPPPLTCQRYC 34
        |||||||
DB      8 csqneyfdisllhacipcqlrcssnppilctqyc 41

RESULT      6
AAE00506    AAE00506 standard; Protein; 184 AA.
ID          AAE00506
IC          AAE00506;

```

XX		31-JUL-2001	(first entry)
DT			
XX			
DE			
XX			
KW	Human B cell maturation protein (BCMA).		
KM			
KX			
XX			
KW	Human: A Proliferation Inducing Ligand Receptor; APRIL-R; cytostatic;		
KM	gene therapy; cancer; nephrotropic; renal disorder; autoimmune disease;		
KX	carcinoma; lung; colon; breast; prostate; Grave's disease; hypertension;		
KW	systemic lupus erythematosus; SLE; inflammation; cardiovascular disease;		
KM	B-cell lympho-proliferative disorder; BCM; immunosuppressive disease;		
KX	organ transplantation; HIV; human immunodeficiency virus; TNF;		
KW	tumour necrosis factor; BCMA; B cell maturation protein.		
KX			
OS	Homo sapiens.		
XX			
PN	WO200124811-A1.		
XX			
PD	12-APR-2001.		
XX			
PF	05-OCT-2000; 2000WO-US27579.		
PR	06-OCT-1999; 99US-0157933.		
XX	11-FEB-2000; 2000US-0181807.		
PR	30-JUN-2000; 2000US-0215688.		
XX			
PA	(BIOJ ) BIOGEN INC.		
XX	(APOT-) APOTEC R & D SA.		
PI	Schneider F, Thompson J, Cachero T, Ambrose C, Rennett P;		
XX			
DR	WPI: 2001-266242/27.		
N-PSDB:	AAD03844.		
XX			
PT	Treating a mammal for a condition associated with undesired cell		
PT	proliferation such as cancer or carcinoma, comprises administering a		
PT	composition comprising A proliferation inducing ligand receptor		
PT	(APRIL-R) antagonist -		
XX			
PS	Claim 3; Fig 3A; 85pp; English.		
XX			
CC	The invention relates to a method of treating a mammal for a condition		
CC	associated with undesired cell proliferation such as cancer or		
CC	carcinoma. The method involves administering a composition comprising		
CC	A proliferation inducing ligand receptor (APRIL-R) also referred as		
CC	B cell maturation protein (BCM or BCMA) antagonist that antagonises the		
CC	interaction between APRIL and its cognate receptor(s). This method is		
CC	useful for treating undesired cell proliferation such as cancer or		
CC	carcinoma e.g. human lung carcinoma, colon carcinoma, breast carcinoma,		
CC	prostate carcinoma, and other carcinomas whose proliferation is modulated		
CC	by APRIL. It is also useful for treating autoimmune diseases (Grave's		
CC	disease, systemic lupus erythematosus-SLE); hypertension, cardiovascular		
CC	diseases, renal disorders, B-cell lympho-proliferative disorders,		
CC	immunosuppressive diseases, organ transplantation, inflammation and		
CC	human immunodeficiency virus (HIV), and for treating, suppressing or		
CC	altering an immune response involving a signalling pathway between		
CC	APRIL-R and its ligand. APRIL-R DNA is also useful in gene therapy.		
CC	The present sequence is human APRIL-R also referred as BCMA or		
CC	BCM protein.		
XX			
SO	Sequence 184 AA;		
Query Match	40.4%; Score 201; DB 22; Length 184;		
Best Local Similarity	100.0%; Pred. NO. 2.8e-10;		
Matches 34; Conservative	0; Mismatches 0; Indels 0; Gaps 0;		
QY	1 CSONEYFDLHACIPCOLRCSSNPPLTCORVC 34		
Db	8 csneyfdisllhacipqqlrcssntpptltqrvc 41		
RESULT	7		
LAB60698			

ID AAB60698 standard; Protein; 184 AA.  
XX AAB60698;  
AC  
XX  
XX  
DT 22-MAY-2001 (first entry)  
XX  
XX  
DE Human BAFR receptor (BAFR-R).  
XX  
KM Human BAFR-R; BAFR receptor; TNF family; immunoregulatory agent;  
KM immune-related disorder; B-cell growth inhibitor; BCMA;  
KM B-cell maturation inhibitor; immunoglobulin production inhibitor;  
KM autoimmune disorder; B-cell lymphoproliferative disorder; hypertension;  
KM renal disorder; immunosuppressive disorder; HIV infection;  
KM organ transplantation; anti-inflammatory; systemic lupus erythematosus;  
KM autoimmune haemolytic anaemia; Grave's disease; multiple myeloma;  
KM B-cell carcinoma; leukaemia; rapidly progressive glomerulonephritis;  
KM lymphoma; gene therapy; cancer; tumour.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200112812-A2.  
PN  
XX  
XX 22-FEB-2001.  
PD  
XX  
XX 16-AUG-2000; 2000WO-US22507.  
PF  
XX  
XX 17-AUG-1999; 99US-0149378.  
PR  
XX 11-FEB-2000; 2000US-0181684.  
PR 18-FEB-2000; 2000US-0183536.  
XX  
XX  
XX (BIOJ) BIOGEN INC.  
PA (APOT-) APOTEC R & D SA.  
XX  
PI Mackay F, Brownlind J, Ambrose C, Tschopp J, Schneider P;  
PI Thompson J;  
XX  
XX WPI: 2001-202866/20.  
DR N-PSDB: AAF59998.  
XX  
XX  
PT Inhibiting dendritic cell-induced B-cell growth, maturation and B-cell  
PT lympho-proliferative disorder by administering BAFR-receptor  
PT polypeptide, chimeric molecule comprising receptor or anti-BAFR-R  
PT antibody homolog  
XX  
XX  
PS Claim 20; Fig 1; 59pp; English.  
XX  
XX The invention relates to the use of a BAFR receptor (BAFR-R, also known  
CC as BCMA) protein, or a BAFR-R fusion protein as an agent for the  
CC treatment of a variety of immune-related disorders. BAFR-R is a member of  
CC the TNF (tumour necrosis factor) family, acting as an immunoregulatory  
CC agent, and also plays a role in the development of hypertension and  
CC related disorders. BAFR-R, fusion proteins containing it, and BAFR-R-  
CC specific antibodies can be used for inhibiting B-cell growth, dendritic  
CC cell-induced B-cell growth and maturation, and immunoglobulin production,  
CC and in the treatment of autoimmune disorders, B-cell lymphoproliferative  
CC disorders, hypertension and renal disorders. The BAFR-R proteins may also  
CC be used in the treatment of immunosuppressive disorders and HIV  
CC infection, and in patients undergoing organ transplantation. The BAFR-R  
CC proteins or BAFR-R specific antibodies may be used for treating,  
CC suppressing or altering an immune response involving a signalling pathway  
CC between BAFR-R and BAFR, thereby inhibiting inflammation. Since BAFR-R  
CC inhibits B-cell growth and maturation it is useful for treating diseases  
CC such as systemic lupus erythematosus, autoimmune haemolytic anaemia,  
CC Grave's disease, multiple myeloma, B-cell carcinomas, leukaemia, rapidly  
CC progressive glomerulonephritis, and lymphomas. Nucleic acids encoding  
CC human BAFR-R may be used in gene therapy to treat tumours, lymphomas,  
CC autoimmune disorders and inherited B-cell-associated disorders. The  
CC present sequence represents human BAFR-R.  
XX  
XX Sequence 184 AA;  
SQ

Query Match

40.4%; Score 201; DB 22; Length 184;

Best Local Similarity 100.0%; Pred. No. 2.8e-10;  
Matches 34; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 CSQNEFYDSLHACIPQQLRCSSNTPPLTCQRYC 34  
DB 8 csqneyfdsllhacipqqlrcssntppltcqryc 41  
RESULT 8  
AAY71979  
ID AAY71979 standard; Protein; 184 AA.  
XX  
XX AAY71979;  
AC  
XX  
DT 28-MAR-2001 (first entry)  
XX  
XX Human B cell maturation factor (BCMA) protein.  
DE  
XX  
XX Human; Tumour Necrosis Factor; TNF; immunosuppressant; TALL-1;  
KM Tumour necrosis factor and Apol-related leucocyte expressed ligand 1;  
KM therapy; autoimmune disorder; rheumatoid arthritis; multiple sclerosis;  
KM systemic lupus erythematosus; SLE; insulin dependent diabetes mellitus;  
KM thrombocytopenia purpura; acute rheumatic fever; Goodpasture's syndrome;  
KM haemolytic anaemia; Grave's disease; myasthenia gravis; chromosome 16;  
KM post-streptococcal glomerulonephritis; polyarteritis nodosa; BCMA;  
KM B cell maturation factor; pemphigus vulgaris; B-lymphocyte proliferation.  
XX  
XX Homo sapiens.  
OS  
XX  
XX  
XX Key Location/Qualifiers  
FH 1..62  
FT Domain  
FT /label= Extracellular\_domain  
XX  
XX WO2000068378-A1.  
XX  
XX 16-NOV-2000.  
PD  
XX  
XX 05-MAY-2000; 2000WO-US12266.  
PF  
XX  
XX 06-MAY-1999; 99US-0132892.  
PR 01-MAY-2000; 2000US-0201012.  
XX  
XX (NAJE-) NAT JEWISH MEDICAL & RES CENT.  
XX  
XX Shu HS;  
PI  
XX  
XX WPI: 2001-016094/02.  
DR N-PSDB: AAD02125.  
XX  
XX  
PT Isolated TALL-1 protein is used to identify compounds that regulate B  
PT lymphocyte proliferation, used to treat B lymphocyte associated  
PT autoimmune disorders -  
XX  
XX  
PS Claim 37; Page 104-105; 112pp; English.  
XX  
XX The present invention relates to Tumour necrosis factor (TNF) and  
CC Apol-related leucocyte expressed ligand 1 (TALL-1) nucleic acid  
CC molecules, proteins (including homologues), and their antibodies. The  
CC invention in particular relates to methods for regulating the  
CC interaction between TALL-1 and TALL-1 receptors (BCMA referred as B cell  
CC maturation factor) to regulate monocyte, macrophage and B lymphocyte  
CC mediated immune responses. TALL-1 protein is useful for identifying  
CC compounds that regulate B lymphocyte proliferation. It is also useful for  
CC treating B lymphocyte associated autoimmune disorders like rheumatoid  
CC arthritis, systemic lupus erythematosus (SLE), insulin dependent diabetes  
CC mellitus, multiple sclerosis, myasthenia gravis, Grave's disease,  
CC autoimmune haemolytic anaemia, autoimmune thrombocytopenia purpura,  
CC Goodpasture's syndrome, pemphigus vulgaris, acute rheumatic fever,  
CC post-streptococcal glomerulonephritis, or polyarteritis nodosa.  
CC The TALL-1 protein and its corresponding nucleic acid sequence are also  
CC useful in diagnostic assays.  
CC The present sequence is a human B cell maturation factor (BCMA)  
CC protein. It is the receptor for TALL-1 protein. BCMA gene is







PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
PS Claim 20; SEQ ID No 38696; 103pp; English.  
XX  
CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 256 AA:

Query Match 19.1%; Score 95; DB 22; Length 256;  
Best Local Similarity 30.6%; Pred. No. 0.49;  
Matches 22; Conservative 7; Mismatches 19; Indels 24; Gaps 4;

OY 16 PCOLRCSSNTPPLTCORCCERYDLSLHAC-----PCLR--CSPTTCOYCFHSEYFDSL 69  
DB 101 pcqsgtscetpcscqscq-----pacttspcqqacvpcvckpvc----- 144  
OY 70 HACPPATCOPYC 81  
DB 145 --cvpvcckpvc 154

RESULT 15  
ABG23698  
ID ABG23698 standard; Protein; 351 AA.  
XX  
AC ABG23698;  
XX  
DT 18-FEB-2002 (first entry)  
XX  
DE Novel human diagnostic protein #23689.  
XX  
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;  
KW food supplement; medical imaging; diagnostic; genetic disorder.  
XX  
OS Homo sapiens.  
XX  
PN WO200175067-A2.  
XX  
PD 11-OCT-2001.  
XX  
PF 30-MAR-2001; 2001WO-US08631.  
XX  
PR 31-MAR-2000; 2000US-0540217.  
PR 23-AUG-2000; 2000US-0649167.  
XX  
PA (HYSE-) HYSEQ INC.  
XX  
PI Drmanac RT, Liu C, Tang YT;  
XX  
DR WPI; 2001-639362/73.

DR N-PsDB; AAS87885.

XX  
PT New isolated polynucleotide and encoded polypeptides, useful in  
PT diagnostics, forensics, gene mapping, identification of mutations  
PT responsible for genetic disorders or other traits and to assess  
PT biodiversity  
PS Claim 20; SEQ ID No 54057; 103pp; English.  
XX  
XX

CC The invention relates to isolated polynucleotide (I) and  
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,  
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome  
CC and gene mapping and in recombinant production of (II). The  
CC polynucleotides are also used in diagnostics as expressed sequence tags  
CC for identifying expressed genes. (I) is useful in gene therapy techniques  
CC to restore normal activity of (II) or to treat disease states involving  
CC (II). (II) is useful for generating antibodies against it, detecting or  
CC quantitating a polypeptide in tissue, as molecular weight markers and as  
CC a food supplement. (II) and its binding partners are useful in medical  
CC imaging of sites expressing (II). (I) and (II) are useful for treating  
CC disorders involving aberrant protein expression or biological activity.  
CC The polypeptide and polynucleotide sequences have applications in  
CC diagnostics, forensics, gene mapping, identification of mutations  
CC and to produce other types of data and products dependent on DNA and  
CC amino acid sequences. ABG00010-ABG30377 represent novel human  
CC diagnostic amino acid sequences of the invention.  
CC Note: The sequence data for this patent did not appear in the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pcl\_sequences.  
XX  
SQ Sequence 351 AA:

Query Match 19.1%; Score 95; DB 22; Length 351;  
Best Local Similarity 30.6%; Pred. No. 0.64;  
Matches 22; Conservative 7; Mismatches 19; Indels 24; Gaps 4;

OY 16 PCOLRCSSNTPPLTCORCCERYDLSLHAC-----PCLR--CSPTTCOYCFHSEYFDSL 69  
DB 196 pcqsgtscetpcscqscq-----pacttspcqqacvpcvckpvc----- 229  
OY 70 HACPPATCOPYC 81  
DB 240 --cvpvcckpvc 249

Search completed: May 15, 2002, 09:49:02  
Job time: 47 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 15, 2002, 09:48:20 ; Search time 14.98 seconds  
(without alignments)  
132.074 Million cell updates/sec

Title: US-09-854-864-13

Perfect score: 498

Sequence: 1 CSQNEYPDSLHACIPQCLR.....SEYFDSLHACIPATCPYC 81

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issuel\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCRTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	93	18.7	166	2	US-08-810-572A-6
2	93	18.7	166	4	US-09-290-333-6
3	93	18.7	293	2	US-08-810-572A-2
4	93	18.7	293	4	US-09-290-333-2
5	82	16.5	508	4	US-09-019-095A-8
6	82	16.5	521	4	US-09-019-095A-22
7	82	16.5	526	4	US-09-019-095A-2
8	80	16.1	3075	2	US-08-460-309-5
9	80	16.1	3075	2	US-08-125-077-5
10	80	16.1	5405	4	US-08-718-388-9
11	79	15.9	545	4	US-09-019-095A-38
12	78	15.7	109	2	US-08-527-044-2
13	78	15.7	109	3	US-09-013-780-2
14	77.5	15.6	139	3	US-08-965-903B-20
15	76.5	15.4	2414	1	US-08-227-536-2
16	76.5	15.4	2414	1	US-08-227-536-2
17	76	15.3	2441	1	US-08-194-468-2
18	76	15.3	2441	3	US-08-961-732-2
19	75.5	15.2	156	3	US-08-600-982-30
20	75.5	15.2	156	5	US-08-600-982-30
21	75.5	15.2	1713	3	US-08-600-982-24
22	75.5	15.2	1713	3	US-08-600-982-24
23	75	15.1	219	2	US-08-152-019A-24
24	75	15.1	219	2	US-08-152-019A-31
25	75	15.1	219	2	US-08-460-309-18
26	75	15.1	430	3	US-08-125-077-18
27	75	15.1	430	3	US-08-997-897-2
			430	4	US-09-156-836B-2

28	73	14.7	689	4	US-09-177-249-2	Sequence 2, Appl
29	73	14.7	689	4	US-09-061-769A-2	Sequence 2, Appl
30	72.5	14.6	969	2	US-08-284-941-2	Sequence 2, Appl
31	72.5	14.6	969	2	US-08-447-642-2	Sequence 2, Appl
32	72.5	14.6	969	4	US-09-236-503-2	Sequence 2, Appl
33	72.5	14.6	969	5	US-09-236-503-2	Sequence 2, Appl
34	72	14.5	275	1	US-08-312-870-7	Sequence 7, Appl
35	72	14.5	366	4	US-08-857-076-103	Sequence 103, App
36	72	14.5	446	1	US-08-307-444A-5	Sequence 5, Appl
37	72	14.5	446	1	US-08-587-389-5	Sequence 5, Appl
38	72	14.5	456	1	US-08-307-444A-3	Sequence 3, Appl
39	72	14.5	456	1	US-08-307-444A-4	Sequence 4, Appl
40	72	14.5	456	1	US-08-587-389-3	Sequence 3, Appl
41	72	14.5	456	1	US-08-587-389-4	Sequence 4, Appl
42	72	14.5	475	1	US-08-307-444A-1	Sequence 1, Appl
43	72	14.5	475	1	US-08-307-444A-2	Sequence 2, Appl
44	72	14.5	475	1	US-08-587-389-1	Sequence 1, Appl
45	72	14.5	475	1	US-08-587-389-2	Sequence 2, Appl

## ALIGNMENTS

RESULT 1  
US-08-810-572A-6  
; Sequence 6, Application US/08810572A  
; Patent No. 5969102

## GENERAL INFORMATION:

APPLICANT: Bram, Richard J.

TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS

TITLE OF INVENTION: CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF USE

TITLE OF INVENTION: THEREOF

NUMBER OF SEQUENCES: 11

CORRESPONDENCE ADDRESS:

ADDRESS: David A. Jackson, Esq.

STREET: 411 Hackensack Ave, Continental Plaza, 4th

FLOOR

CITY: Hackensack

STATE: New Jersey

COUNTRY: USA

ZIP: 07601

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/810,572A

FILING DATE: 28-FEB-1997

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Jackson Esq., David A.

REGISTRATION NUMBER: 26,742

TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800

TELEFAX: 201-343-1684

INFORMATION FOR SEQ. ID NO. 6:

SEQUENCE CHARACTERISTICS:

LENGTH: 166 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

HYPOTHETICAL: NO

FRAGMENT TYPE: N-terminal

ORIGINAL SOURCE:

ORGANISM: Homo sapiens

US-08-810-572A-6

Query Match 18.7%; Score 93; DB 2; Length 166;

US-08-810-572A-2  
: Sequence 2, Application US/08810572A  
: Patent No. 5969102  
: GENBRAT INFORMATION.

APPLICANT: Bram, Richard J.  
APPLICANT: von Bulow, Gotz  
TITLE OF INVENTION: A LYMPHOCYTE SURFACE RECEPTOR THAT BINDS

NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:

STREET: 411 Hackensack Ave, Continental Plaza, 4th  
FLOOR  
CITY: Hackensack  
STATE: NJ  
COUNTRY: USA

```

;
;      COUNTRY:  USA
;      ZIP:      07601
;
;  COMPUTER READABLE FORM:  31-1-

```

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/810,572A  
 FILING DATE: 28-FEB-1997  
 CLASSIFICATION: 536  
 AMENDMENT/ACR/INT: INFOBULLETIN

NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 1340-1-007  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEO ID NO: 2:  
SEQUENCE CHARACTERISTICS.

LENGTH: 293 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

?  
? MOLECULE TYPE: Protein  
? HYPOTHETICAL: NO  
? FRAGMENT TYPE: N-terminal  
ORIGINAL SOURCE:

US-08-810-572A-2

Query match	18.7%	Score 93	DB 2	Length 293
Best Local Similarity	28.4%	Pred. 0.057		
Matches	21	Conservative	13	Mismatches 22
				Indels 18
				Gaps 5

34 CPBEOYMDPLGLGTCSMCKTKTICNHQS-QRTCAAFCSRLSCREKGRFYDHLIR--DCISCA 90

QY	32	-----PRICQRC	59
Db	91	SICGHPKOCAYFC	104

RESULT 4  
US-09-290-333-2  
Sequence 2. Application US/09290333

AGENT NO. 0310ZZ  
GENERAL INFORMATION:  
APPLICANT: Bram, Richard J.  
von Bulow, Goltz

LITTLE OF INVENTION: A LIPIDIC SURFACE RECEPTOR THAT BINDS  
 CAML, NUCLEIC ACIDS ENCODING THE SAME AND METHODS OF



SEQ ID NO 2  
LENGTH: 526  
TYPE: PRT  
ORGANISM: Murine  
US-09-019-095A-2

Query Match 16.5%; Score 82; DB 4; Length 526;  
Best Local Similarity 35.8%; Pred. No. 1.1;  
Matches 19; Conservative 8; Mismatches 18; Indels 8; Gaps 2;

QY 19 LRCSNTPPLTCQRCCEYFDSLHACPLRCSPTQYCCPHSEYFDSLHA 71  
Db 66 LQCLTHRPPL-----ADYMLSOEHSQTC--CSPEGCKLCAMEALVTSLSLHS 110

RESULT 8  
US-08-460-309-5  
Sequence 5, Application US/08460309  
Patent No. 5837496  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leivo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/460.309  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/125.077  
FILING DATE: 22-SEP-1993  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472.319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919.951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3075 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-460-309-5

Query Match 16.1%; Score 80; DB 2; Length 3075;  
Best Local Similarity 23.4%; Pred. No. 9;  
Matches 29; Conservative 10; Mismatches 37; Indels 48; Gaps 8;  
QY 1 CSQNEYFDSL--HACIPQ-----LRCSNTPPLTCQRCCE-- 36

Db 831 CADGYGNPTVPGESCVCDSGNVDPSEAGHCDSTGECIKLGNTDGAHCR-CADGF 889  
QY 37 YFDSL-----LHACPC-----LRCSNTPPLTCQRCCE-- 73  
Db 890 YGDVATKNCRCRCHVKGSHAVCHLETGLCDCKPNTVGQCDQCLHGYGLDGHGCR 949  
QY 74 PATC 77  
Db 950 PCNC 953

RESULT 9  
US-08-125-077-5  
Sequence 5, Application US/08125077  
Patent No. 5872231  
Patent No. 5872231 5840863  
GENERAL INFORMATION:  
APPLICANT: Engvall, Eva  
APPLICANT: Leivo, Ilmo  
TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
TITLE OF INVENTION: Fragments and Uses Thereof  
NUMBER OF SEQUENCES: 23  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Campbell and Flores  
STREET: 4370 La Jolla Village Drive, suite 700  
CITY: San Diego  
STATE: California  
COUNTRY: USA  
ZIP: 92122  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/125.077  
FILING DATE: 22-SEP-1993  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US PCT/US 94/10730  
FILING DATE: 21-SEP-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/472.319  
FILING DATE: 30-JAN-1990  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/919.951  
FILING DATE: 27-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Campbell, Cathryn A.  
REGISTRATION NUMBER: 31,815  
REFERENCE/DOCKET NUMBER: P-LA 9721  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 535-9001  
TELEFAX: (619) 535-8949  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3075 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-125-077-5

Query Match 16.1%; Score 80; DB 2; Length 3075;  
Best Local Similarity 23.4%; Pred. No. 9;  
Matches 29; Conservative 10; Mismatches 37; Indels 48; Gaps 8;

QY 1 CSQNEYFDSL--HACIPQ-----LRCSNTPPLTCQRCCE-- 36  
Db 831 CADGYGNPTVPGESCVCDSGNVDPSEAGHCDSTGECIKLGNTDGAHCR-CADGF 889  
QY 37 YFDSL-----LHACPC-----LRCSNTPPLTCQRCCE-- 73

Db 89 YGDAVTAKNCRACECHVKSHSAVCHLETGLCDCKBNVTGQDCDOCLHGYGLDSHGCR 949  
OY 74 PATC 77  
Db 950 PCNC 953

RESULT 10  
US-08-718-388-9  
; Sequence 9, Application US/08718388  
; Patent No. 6271362  
; GENERAL INFORMATION:  
; APPLICANT: MORIKAWA, MINORU  
; APPLICANT: HARADA, NAOKI  
; TITLE OF INVENTION: GENE ENCODING IGG FC REGION-BINDING  
; TITLE OF INVENTION: PROTEIN  
; NUMBER OF SEQUENCES: 29  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH  
; STREET: PO BOX 747  
; CITY: FALLS CHURCH  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22040-0747  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/718,388  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: MURPHY JR, GERALD M  
; REGISTRATION NUMBER: 28,977  
; REFERENCE/DOCKET NUMBER: 0230-111  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703) 205-8000  
; TELEFAX: (703) 205-8050  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5405 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-718-388-9

Query Match 16.1%; Score 80; DB 4; Length 5405;  
Best Local Similarity 25.4%; Pred. No. 15;  
Matches 29; Conservative 12; Mismatches 25; Indels 48; Gaps 10;

OY 1 CGSNEFYDSLHRCI-PCQARCSNTPPLTCQRYCCEYFSLHACPC-----LRCS 51  
Db 2733 CPONSIVE-----CADTCSLGSALSAPLQCPDCAE-----GQCDSGFLYNGQACV 2781

OY 52 PPTCQYRC-CFHS-EYFD---SLF-----HACPPA-TCOP 79  
Db 2782 P--IQCCGCHNAYEPEQTVLIDMCROOCTCHAGKYVVCQHSCKPQVCP 2833

RESULT 11  
US-09-019-095A-38  
; Sequence 38, Application US/09019095A  
; Patent No. 6287858  
; GENERAL INFORMATION:  
; APPLICANT: D'Andrea, Alan D.  
; APPLICANT: Zhu, Yuan  
; TITLE OF INVENTION: Deubiquitinating Enzymes That Regulate  
; TITLE OF INVENTION: Cell Growth  
; FILE REFERENCE: DEPT-4352A2  
; CURRENT APPLICATION NUMBER: US/09/019,095A

; CURRENT FILING DATE: 1998-02-05  
; PRIOR APPLICATION NUMBER: PCF/US96/12834  
; PRIOR FILING DATE: 1996-08-07  
; PRIOR APPLICATION NUMBER: US 60/002,066  
; PRIOR FILING DATE: 1995-08-09  
; PRIOR APPLICATION NUMBER: US 60/019,787  
; PRIOR FILING DATE: 1996-06-14  
; NUMBER OF SEQ ID NOS: 51  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 38  
; LENGTH: 545  
; TYPE: PRT  
; ORGANISM: murine  
US-09-019-095A-38

Query Match 15.9%; Score 79; DB 4; Length 545;  
Best Local Similarity 34.0%; Pred. No. 2.1;  
Matches 18; Conservative 9; Mismatches 18; Indels 8; Gaps 2;

OY 19 LRCSNTPPLTCQRYCCEYFSLHACPCRLCSPTTCQYCCFHSYFDSLHA 71  
Db 66 LQCLTHPL-----ADYMLSOBYSTQC--CSPEGCKMCAEHWYQSILHS 110

RESULT 12  
US-08-527-044-2  
; Sequence 2, Application US/08527044  
; Patent No. 5885568  
; GENERAL INFORMATION:  
; APPLICANT: BUNSTEAD, Janene Marilyn J.M.  
; APPLICANT: TOMLEY, Fiona Margaret F.  
; APPLICANT: DUNN, Patrick Paul James P.  
; APPLICANT: VERMEULEN, Arnoldus Nicolaas A.N.  
; TITLE OF INVENTION: Coccidiosis poultry vaccine  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Akzo No. 5885568e1 N.V  
; STREET: 1300 PICCARD DRIVE #206  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850-4373  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/527,044  
; FILING DATE: 12-SEP-1995  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: EP 9420267.6  
; FILING DATE: 16-SEP-1994  
; ATTORNEY/AGENT INFORMATION:  
; NAME: GORMLEY, MARY E.  
; REGISTRATION NUMBER: 34,409  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 109 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-527-044-2

Query Match 15.7%; Score 78; DB 2; Length 109;  
Best Local Similarity 26.7%; Pred. No. 0.58;  
Matches 20; Conservative 7; Mismatches 24; Indels 24; Gaps 5;

OY 10 LHAICPQARCSNTPPLTCQRYCCEYFSLHACPCRLCSPTTC---QYCCFHSYFD 66  
Db 6 LTNAC--CCLRYTNS-----CCSKYCS-----KCCCSKCCSKCCCTYCC----- 45



CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/227,536  
FILING DATE: 14-APR-1994  
CLASSIFICATION: 436  
ATTORNEY/AGENT INFORMATION:  
NAME: William Ph.D., Kathleen A.  
REGISTRATION NUMBER: 34,380  
REFERENCE/DOCKET NUMBER: DFCI-308XX  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-2290  
TELEFAX: (617) 451-0313  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2414 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: Protein  
US-08-227-536-2

Query Match 15.4%; Score 76.5; DB 1; Length 2414;  
Best Local Similarity 21.7%; Pred. No. 15;  
Matches 30; Conservative 15; Mismatches 22; Indels 71; Gaps 9;  
QY 1 CSQNEYFDSLHACIPC-----QLRCSSNTP-----PLTCORY 33  
|:|:|  
|:|:|  
DB 1683 CTVCEDYD-----LCITCYNTKKNHDKMKELGLDDESNNOQAATQSPGDSRRLSIQR- 1737  
|:|:|  
|:|:|  
QY 34 CCEYFDSLHACPC--LRCSPTCQ-----YCCPHSE 63  
|:|:|  
|:|:|  
DB 1738 CIO---SLVHACQCRNANCSLPSCQMKRVVQHTKGCKRRTNGGCPICKOLIALCCYHAK 1794  
|:|:|  
|:|:|  
QY 64 YFDSLHACPPATCOPYC 81  
|:|  
|:|  
DB 1795 HCOE--NKCP---VPFC 1806  
|:|  
|:|

Search completed: May 15, 2002, 09:50:03  
Job time: 103 sec

